

Hiroyuki Yoshida
David Hawkes
Michael W. Vannier (Eds.)

LNCS 7601

Abdominal Imaging

Computational and Clinical Applications

4th International Workshop
Held in Conjunction with MICCAI 2012
Nice, France, October 2012, Proceedings

 **Springer**

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

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

Hiroyuki Yoshida David Hawkes
Michael W. Vannier (Eds.)

Abdominal Imaging

Computational and Clinical Applications

4th International Workshop
Held in Conjunction with MICCAI 2012
Nice, France, October 1, 2012
Proceedings

 Springer

Volume Editors

Hiroyuki Yoshida
General Hospital and Harvard Medical School
3D Imaging Research, Department of Radiology
25 New Chardon Street, Suite 400C
Boston, MA 02114, USA
E-mail: yoshida.hiro@mgh.harvard.edu

David Hawkes
University College London
Centre for Medical Image Computing
Malet Place
London, WC1E 6BT, UK
E-mail: d.hawkes@cs.ucl.ac.uk

Michael W. Vannier
The University of Chicago
Department of Radiology
5841 South Maryland Avenue, MC2026
Chicago, IL 60637, USA
E-mail: mvanier@radiology.bsd.uchicago.edu

ISSN 0302-9743
ISBN 978-3-642-33611-9
DOI 10.1007/978-3-642-33612-6
Springer Heidelberg Dordrecht London New York

e-ISSN 1611-3349
e-ISBN 978-3-642-33612-6

Library of Congress Control Number: 2012947419

CR Subject Classification (1998): I.4.9, I.4.1, I.4.3, I.4.6-9, J.3, F.1.1

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

© Springer-Verlag Berlin Heidelberg 2012

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.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

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

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The Fourth International Workshop on Computational and Clinical Applications in Abdominal Imaging was held in conjunction with the 15th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) on October 1, 2012, in Nice, France.

In the abdomen, organs and disease appearance are complex and subtle, and thus the development of computational models that are useful in clinical practice is highly challenging. Nevertheless, diagnosis often relies on the quantitative measures of organs and lesions, because their volumes and shapes are strong indicators of disorders. Given the complexity and high variability of abdominal organs, the identification of distinct computational challenges for integrative models of organs and abnormalities is essential for understanding anatomy and disease, evaluating treatment, and planning intervention.

Leveraging the success of the previous workshops, the fourth MICCAI workshop on computational abdominal imaging aimed to provide a comprehensive forum for reviewing clinical opportunities in computational abdominal imaging, and for sharing state-of-the-art as well as emerging techniques for solving computationally challenging image analysis and visualization problems, by bringing together leading researchers and clinician-scientists from around the world.

In response to a call for papers, a total of 37 papers were initially submitted to the workshop. These papers underwent a rigorous, double-blind peer-review process, with each paper being reviewed by a minimum of 2, and in most cases, by 3 expert reviewers from the Scientific Review Committee. Based on the results of this review, 31 papers were accepted by the workshop. All of the accepted papers were revised by incorporating the reviewers' comments and re-submitted by the authors to be included in this proceedings volume.

The workshop provided participants with in-depth and interactive discussions on the emerging techniques introduced by the current clinical challenges that were not fully explored in the MICCAI main conference. The workshop successfully provided a forum among participants for the dissemination of state-of-the-art research and technologies, the exchange of emerging ideas, the initiation of collaborations, and the exploration of new clinical applications for diagnostic and interventional procedures in abdominal imaging.

We would like to express our sincere appreciation to the authors whose contributions to this proceedings book have required considerable commitment of time and effort. We also thank the members of the Scientific Review Committee for their excellent work in reviewing the submitted manuscripts on a tight schedule.

October 2012

Hiroyuki Yoshida
David Hawkes
Michael Vannier

Workshop Organization

Organizing Committee

Hiroyuki Yoshida	Massachusetts General Hospital/Harvard Medical School, USA
David Hawkes	University College London, UK
Michael Vannier	The University of Chicago, USA
Nicholas Ayache	Inria, France

Workshop Committee

Wenli Cai	Massachusetts General Hospital/Harvard Medical School, USA
Philippe C. Cattin	Medical Image Analysis Center/University of Basel, Switzerland
Aly A. Farag	University of Louisville, USA
Hiroshi Fujita	Gifu University, Japan
Lubomir Hadjiyski	University of Michigan, USA
Kenneth R. Hoffmann	University at Buffalo, USA
David R. Holmes	Mayo Clinic College of Medicine, USA
Mingxing Hu	University College London, UK
Lakhmi C. Jain	University of South Australia, Australia
Jong Hyo Kim	Seoul National University, South Korea
Se Hyung Kim	Seoul National University, South Korea
Marius George Linguraru	Children's National Medical Center, USA
Honbing Lu	Fourth Military Medical University, China
Kensaku Mori	Nagoya University, Japan
Yuichi Motai	Virginia Commonwealth University, USA
Sandy Napel	Stanford University, USA
Emanuele Neri	University of Pisa, Italy
Noboru Niki	Tokushima University, Japan
Kazunori Okada	San Francisco State University, USA
Daniele Regge	Institute for Cancer Research and Treatment, Italy
Georgios Sakas	Technische Universität Darmstadt/Fraunhofer Institute for Computer Graphics Research IGD, Germany
Yoshinobu Sato	Osaka University Graduate School of Medicine, USA
Akinobu Shimizu	Tokyo University of Agriculture and Technology, Japan

VIII Workshop Organization

Frans Voss
Delft University of Technology/Academic
Medical Center Amsterdam,
The Netherlands

Simon K. Warfield
Children's Hospital Boston/Harvard
Medical School, USA

Program Chair

Janne Näppi
Massachusetts General Hospital/Harvard
Medical School, USA

Program Committee

Dongqing Chen
Yufei Chen
Klaus Drechsler
Xianfeng Gu
June-Goo Lee
Sang-Ho Lee
Freiman Moti
Masahiro Oda
Danielle Pace
Frank Preiswerk
Mandava Rajeswari
Holger Roth
Sergio Vera
Thomas Wittenberg
Xin Zhao
Hongbin Zhu

GE Healthcare, USA
Tongji University, China
Fraunhofer Institute for Computer Graphics
Research IGD, Germany
State University of New York at Stony Brook,
USA
University of Pittsburgh, USA
Massachusetts General Hospital/Harvard
Medical School, USA
Children's Hospital Boston/Harvard Medical
School, USA
Nagoya University, Japan
Kitware Inc., USA
University of Basel c/o University Hospital
Basel, Switzerland
Universiti Sains Malaysia, Malaysia
University College London, UK
Universitat Autònoma de Barcelona, Spain
Fraunhofer Institute for Integrated Circuits IIS,
Germany
Stony Brook University, USA
State University of New York at Stony Brook,
USA

Scientific Review Committee

Wenli Cai
Massachusetts General Hospital/Harvard
Medical School, USA

Philippe C. Cattin
Medical Image Analysis Center/University
of Basel, Switzerland

Dongqing Chen
Yufei Chen
GE Healthcare, USA
Tongji University, China

Klaus Drechsler	Fraunhofer Institute for Computer Graphics Research IGD, Germany
Hiroshi Fujita	Gifu University, Japan
Lubomir Hadjiyski	University of Michigan, USA
Kenneth R. Hoffmann	University at Buffalo, USA
David R. Holmes	Mayo Clinic College of Medicine, USA
Mingxing Hu	University College London, UK
Jong Hyo Kim	Seoul National University, South Korea
Marius George Linguraru	Children's National Medical Center, USA
Xianfeng Gu	State University of New York at Stony Brook, USA
June-Goo Lee	University of Pittsburgh, USA
Sang-Ho Lee	Massachusetts General Hospital/Harvard Medical School, USA
Honbing Lu	Fourth Military Medical University, China
Kensaku Mori	Nagoya University, Japan
Freiman Moti	Children's Hospital Boston/Harvard Medical School, USA
Yuichi Motai	Virginia Commonwealth University, USA
Sandy Napel	Stanford University, USA
Noboru Niki	Tokushima University, Japan
Masahiro Oda	Nagoya University, Japan
Kazunori Okada	San Francisco State University, USA
Danielle Pace	Kitware Inc., USA
Frank Preiswerk	University of Basel c/o University Hospital Basel, Switzerland
Holger Roth	University College London, UK
Georgios Sakas	Technische Universität Darmstadt/Fraunhofer Institute for Computer Graphics Research IGD, Germany
Yoshinobu Sato	Osaka University Graduate School of Medicine, USA
Akinobu Shimizu	Tokyo University of Agriculture and Technology, Japan
Frans Voss	Delft University of Technology / Academic Medical Center Amsterdam, The Netherlands
Simon K. Warfield	Children's Hospital Boston / Harvard Medical School, USA
Thomas Wittenberg	Fraunhofer Institute for Integrated Circuits IIS, Germany
Sergio Vera	Universitat Autònoma de Barcelona, Spain
Xin Zhao	Stony Brook University, USA

Proceedings Editors

Hiroyuki Yoshida	Massachusetts General Hospital/Harvard Medical School, USA
David Hawkes	University College London, UK
Michael Vannier	The University of Chicago, USA

Proceedings Editorial Board

Janne Näppi	Massachusetts General Hospital/Harvard Medical School, USA
Sang-Ho Lee	Massachusetts General Hospital/Harvard Medical School, USA

Table of Contents

Colon and Other Gastrointestinal Tract

Prone to Supine CT Colonography Registration Using a Landmark and Intensity Composite Method	1
<i>Thomas E. Hampshire, Holger R. Roth, Darren J. Boone, Greg Slabaugh, Steve Halligan, and David J. Hawkes</i>	
External Clinical Validation of Prone and Supine CT Colonography Registration	10
<i>Holger R. Roth, Darren J. Boone, Steve Halligan, Thomas E. Hampshire, Jamie R. McClelland, Mingxing Hu, Shonit Punwani, Stuart Taylor, and David J. Hawkes</i>	
Efficient Topological Cleaning for Visual Colon Surface Flattening	20
<i>Rui Shi, Wei Zeng, Jerome Zhengrong Liang, and Xianfeng David Gu</i>	
Adaptive Volumetric Detection of Lesions for Minimal-Preparation Dual-Energy CT Colonography	30
<i>Janne J. Näppi, Se Hyung Kim, and Hiroyuki Yoshida</i>	
Computer-Aided Detection for Ultra-Low-Dose CT Colonography	40
<i>Janne J. Näppi, Masanori Imuta, Yasuyuki Yamashita, and Hiroyuki Yoshida</i>	
Application of CT Simulation Technique for Virtual Ultra-Low-Dose Trial in CT Colonography	49
<i>Chang Won Kim and Jong Hyo Kim</i>	
Application of CT Simulation Technique for Virtual Fecal Tagging in CTC	58
<i>Zepa Yang, Hyeong-min Jin, and Jong Hyo Kim</i>	
Application of CT Acquisition Parameters as Features in Computer-Aided Detection for CT Colonography	69
<i>Janne J. Näppi, Don Rockey, Daniele Regge, and Hiroyuki Yoshida</i>	
Comparative Performance of State-of-the-Art Classifiers in Computer-Aided Detection for CT Colonography	78
<i>Sang Ho Lee, Janne J. Näppi, and Hiroyuki Yoshida</i>	

Piecewise Structural Diffusion Defined on Shape Index for Noise Reduction in Dual-Energy CT Images	88
<i>Wenli Cai, June-Goo Lee, Da Zhang, Christina Piel, and Hiroyuki Yoshida</i>	
A Supervised Learning Based Approach to Detect Crohn’s Disease in Abdominal MR Volumes	97
<i>Dwarikanath Mahapatra, Peter Schueffler, Jeroen A.W. Tielbeek, Joachim M. Buhmann, and Franciscus M. Vos</i>	
Out-of-Plane Motion Compensation in Cine-MRI	107
<i>Mehmet Yigitsoy, Sonja Kirchhoff, Maximilian F. Reiser, and Nassir Navab</i>	
Real-Time Phase Boundary Detection for Colonoscopy Videos Using Motion Vector Templates	116
<i>Ruwan Nawarathna, JungHwan Oh, Jayantha Muthukudage, Wallapak Tavanapong, Johnny Wong, and Piet C. de Groen</i>	

Liver, Kidney, and Other Organs

Relaxed Conditional Statistical Shape Models and Their Application to Non-contrast Liver Segmentation	126
<i>Sho Tomoshige, Elco Oost, Akinobu Shimizu, Hidefumi Watanabe, Hidefumi Kobatake, and Shigeru Nawano</i>	
A Landmark-Based Primal-Dual Approach for Discontinuity Preserving Registration	137
<i>Silja Kiriyanthan, Ketut Fundana, Tahir Majeed, and Philippe C. Cattin</i>	
Predicting Liver Motion Using Exemplar Models	147
<i>Golnoosh Samei, Christine Tanner, and Gabor Székely</i>	
Semi-automatic Discrimination of Normal Tissue and Liver Cancer Lesions in Contrast Enhanced X-Ray CT-Scans	158
<i>Sanat Upadhyay, Manos Papadakis, Saurabh Jain, Gregory Gladish M.D., Ioannis A. Kakadiaris, and Robert Azencott</i>	
Evaluation of Diffusion Filters for 3D CTA Liver Vessel Enhancement	168
<i>Ha Manh Luu, Adriaan Moelker, Camiel Klink, Adrienne Mendrik, Wiro Niessen, and Theo van Walsum</i>	
3D Shape Analysis for Liver-Gallbladder Anatomical Structure Retrieval	178
<i>Weimin Huang, Wei Xiong, Jiayin Zhou, Jing Zhang, Tao Yang, Jiang Liu, Yi Su, Calvin Lim, Chee Kong Chui, and Stephen Chang</i>	

Tumor Sensitive Matching Flow: An Approach for Ovarian Cancer Metastasis Detection and Segmentation	188
<i>Jianfei Liu, Shijun Wang, Marius G. Linguraru, and Ronald M. Summers</i>	
Fast Segmentation of Abdominal Wall: Application to Sliding Effect Removal for Non-rigid Registration	198
<i>Wenwu Zhu, Stephane Nicolau, Luc Soler, Alexandre Hostettler, Jacques Marescaux, and Yves Rémond</i>	
A Fully Automated Framework for Renal Cortex Segmentation	208
<i>Xinjian Chen, Hemin Zhao, and Jianhua Yao</i>	
An Implicit Inter-subject Shape Driven Image Deformation Model for Prostate Motion Estimation	218
<i>Bartłomiej W. Papież, Bogdan J. Matuszewski, Lik-Kwan Shark, and Christopher Moore</i>	
Segmentation of the Cartilage in the Rib Cage in 3D MRI	229
<i>Yolanda H. Noorda, Lambertus W. Bartels, and Josien P.W. Pluim</i>	
Method for Detecting Enlarged Lymph Nodes from 3D Abdominal CT Images with a Multi-shape and Multi-scale Ellipsoidal Structure Detection Filter	238
<i>Masahiro Oda, Takayuki Kitasaka, Michitaka Fujiwara, Kazunari Misawa, and Kensaku Mori</i>	
Delineation of Liver Tumors from CT Scans Using Spectral Clustering with Out-of-Sample Extension and Multi-windowing	246
<i>Jiayin Zhou, Weimin Huang, Wei Xiong, Wenyu Chen, Sudhakar K. Venkatesh, and Qi Tian</i>	
Evaluation of Medical Image Registration by Using 3D SIFT and Phase-Only Correlation	255
<i>Zisheng Li, Tsuneya Kurihara, Kazuki Matsuzaki, and Toshiyuki Irie</i>	
Optimal Medial Surface Generation for Anatomical Volume Representations	265
<i>Sergio Vera, Miguel A. González, Marius George Linguraru, and Debora Gil</i>	
Registration of Free-Breathing Abdominal 3D Contrast-Enhanced CT	274
<i>Blandine Romain, Véronique Letort, Olivier Lucidarme, Florence d'Alché-Buc, and Laurence Rouet</i>	

Non-newtonian Blood Flow Analysis for the Portal Vein Based on a CT Image	283
<i>Harvey Ho, Adam Bartlett, and Peter Hunter</i>	
Tracer Kinetic Modeling by Morales-Smith Hypothesis in Hepatic Perfusion CT	292
<i>Sang Ho Lee, Wenli Cai, and Hiroyuki Yoshida</i>	
Author Index	303

Prone to Supine CT Colonography Registration Using a Landmark and Intensity Composite Method

Thomas E. Hampshire¹, Holger R. Roth¹, Darren J. Boone², Greg Slabaugh³,
Steve Halligan², and David J. Hawkes¹

¹ Centre for Medical Image Computing,
University College London, London, WC1E 6BT, UK
`thomas.hampshire.09@ucl.ac.uk`

² Centre for Medical Imaging, University College Hospital, London, UK
³ Department of Computing, City University, London, UK

Abstract. Matching corresponding location between prone and supine acquisitions for CT colonography (CTC) is essential to verify the existence of a polyp, which can be a difficult task due to the considerable deformations that will often occur to the colon during repositioning of the patient. This can induce error and increase interpretation time. We propose a novel method to automatically establish correspondence between the two acquisitions. A first step segments a set of haustral folds in each view and determines correspondence via a labelling process using a Markov Random Field (MRF) model. We show how the landmark correspondences can be used to non-rigidly transform a 2D source image derived from a conformal mapping process on the 3D endoluminal surface mesh to achieve full surface correspondence between prone and supine views. This can be used to initialise an intensity-based non-rigid B-spline registration method which further increases the accuracy. We demonstrate a statistically significant improvement over the intensity based non-rigid B-spline registration by using the composite method.

Keywords: CT colonography, image registration.

1 Introduction

A number of methods have been proposed to find correspondence between the prone and supine positions. Centreline-based methods extract and align colonic centrelines by stretching and shrinking based on path geometries [12]. These methods are inherently restricted to achieving a registration along a single dimension and do not give any information about the degree of torsion of the colon wall between views. Anatomical landmarks can be used to help align the two datasets by first identifying a stable set of anatomical features, such as the caecum, rectum and flexures [12,6], but stand-alone they do not provide a fine enough registration between views. Voxel-based methods provide a further means of registration [9]. However, these methods rely to varying extents

upon continuous prone and supine colonic segmentations, free from occlusion by fluid or collapse; a scenario which occurs infrequently in daily practice, despite optimal bowel preparation [10].

Fukano et al. proposed a registration method based on haustral fold matching [1]. A second-order derivative difference filter was used to extract folds; their volume and relative positions along the centreline with respect to a set of locations of high centreline curvature were used to establish correspondence. They reported correct registration of 65.1% of large folds, and 13.3% of small folds; where 9.3% and 32.7% of folds could not be judged.

Zeng et al. combined conformal mapping with feature matching between the prone and supine surfaces [14]. The prone and supine colonic segmentations were mapped onto five rectangle pairs. Correspondences were established using a feature matching method based upon mean curvature. The method relied on accurately determining five matching segments in the prone and supine datasets, which is difficult to achieve and may not be possible in the case of local colonic collapse.

Recently, Hampshire et al. [2] presented a method for generating a set of robust landmark correspondences between the prone and supine CT data using haustral folds. A virtual camera registration is used to create a cost function for matching pairs of folds between the prone and supine acquisitions. A Rotation Minimising Frame (RMF) is swept along the centreline to parametrise the 3D fold position to a 2D vector consisting of centreline distance and angular orientation. Additional fold neighbourhood information in this parametrised space is used to enforce geometric constraints in the form of a pair-wise cost function. The cost functions are incorporated into a MRF model, and a fold labelling assignment is achieved by a Belief Propagation (BP) [13] optimisation process.

Roth et al. [7] provide a full surface registration via a conformal mapping of the prone and supine endoluminal surfaces to 2D cylindrical domains using Ricci flow [3,15], followed by a non-rigid cylindrical intensity based registration using a B-spline method [8] with a sum-of-squared-differences similarity metric based on shape index (SI) [4].

This paper introduces a new composite registration method, first using the sparse positions and displacements of the landmark based registration [2] mapped onto a 2D domain created by performing a conformal mapping using the Ricci flow algorithm [3,15], to construct an underlying function based on multilevel B-splines that can be evaluated at any point to give a transformation from the prone to the supine images. This transformation is further refined by the intensity based registration in [7]. We demonstrate a statistically significant improvement over the previously published methods.

2 B-Spline Approximation

We wish to approximate a smooth function f which relates the (x, y) points in the prone unfolded image, to their (x', y') positions in the supine image over domain $\Omega = (x, y) | 0 \leq x < m, 0 \leq y < n$. To do this we use the set of folds $P = \{(x_c, y_c, v_c)\}$ where (x_c, y_c) is a point in Ω and v_c is the value at (x_c, y_c) .

We define the mapping in terms of two functions: $x' = X(x, y)$ and $y' = Y(x, y)$. As these functions can be derived simultaneously, we use the notation $v = (x', y') = f(x, y)$. To approximate the data P , we use function f as a uniform bicubic B-spline, defined by control lattice Φ overlaid on domain Ω using the method in [5]. We also assume Φ is an $(m+3) \times (n+3)$ lattice, where m and n are the image dimensions defined in lattice control points. We define ϕ_{ij} as the value of ij -th control point on lattice Ω for $i = -1, 0, \dots, m+1$ and $j = -1, 0, \dots, n+1$. We can then define the approximation function f :

$$f(x, y) = \sum_{k=0}^3 \sum_{l=0}^3 B_k(s) B_l(t) \phi_{(i+k)(j+l)}, \quad (1)$$

where $i = \lfloor x \rfloor - 1, j = \lfloor y \rfloor - 1, s = x - \lfloor x \rfloor$, and $t = y - \lfloor y \rfloor$. B_k and B_l are basis functions:

$$\begin{aligned} B_0(t) &= (1-t)^3/6, \\ B_1(t) &= (3t^3 - 6t^2 + 4)/6, \\ B_2(t) &= (-3t^3 + 3t^2 + 3t + 1)/6, \\ B_3(t) &= (t^3)/6, \end{aligned} \quad (2)$$

where $0 \leq t < 1$. For every point in $P = \{(x_c, y_c, v_c)\}$ a different value ϕ_c of each of the control points ϕ_{ij} is defined:

$$\phi_c = \frac{w_c v_c}{\sum_{a=0}^3 \sum_{b=0}^3 w_{ab}^2}, \quad (3)$$

where $w_c = w_{kl} = B_k(s) B_l(t), k = (i+1) - \lfloor x_c \rfloor, l = (j+1) - \lfloor y_c \rfloor, s = x_c - \lfloor x_c \rfloor, t = y_c - \lfloor y_c \rfloor$. Only data points in the 4×4 neighbourhood of each control point are taken into consideration. To choose a value for each ϕ_{ij} from the contributions from each point ϕ_c the error $e(\phi_{ij}) = \sum_c (w_c \phi_{ij} - w_c \phi_c)^2$ is minimised by differentiating $e(\phi_{ij})$ with respect to ϕ_{ij} giving:

$$\phi_{ij} = \frac{\sum_c w_c^2 \phi_c}{\sum_c w_c^2}. \quad (4)$$

To allow for a smooth function over the entire domain and more accurate local deformations, a multilevel B-spline approximation is used to generate a hierarchy of control lattices from coarse to fine. A refinement process is used to reduce the sum of these functions into one B-spline function. For each level of control lattice Φ_k we can derive a finer control lattice Φ'_k such that $F(\Phi'_k) = F(\Phi_k)$. We then derive function f_{k+1} by using control lattice Φ_{k+1} to approximate data $P_k = \{(x_c, y_c, \Delta_{k+1} v_c)\}$, where $\Delta_{k+1} v_c = v_c - \sum_{i=0}^k f_i(x_c, y_c) = \Delta_k v_c - f_k(x_c, y_c)$, and $\Delta_0 v_c = v_c$. Each function serves to remove the residual error from the coarser lattice at each level. We can now define a progressive control lattice $\Psi_{k+1} = \Phi'_k + \Phi_{k+1}$ from the coarsest to finest levels. We apply this technique to the images created by a conformal mapping [7] of the endoluminal surfaces

of prone and supine images onto a rectangular domain. The sparse set of data points $P = \{(x_c, y_c, v_c)\}$ have their positional information $\{(x_c, y_c)\}$ taken from the positions of haustral folds mapped onto the 2D domain, and the vertical and horizontal displaced positions $\{v_c\}$ of the corresponding positions in the supine image. To allow for a pseudo-continuous function over the y -axis, the image is tripled (figure 1). Due to the true cylindrical nature of the registration problem, there is an ambiguity over the direction of vertical displacement in the 2D images. To create a smooth displacement, the B-spline fitting is repeated and at each iteration the datum P_c with the maximum error between the y component of the estimated and true displacement $e_y = |(F(x_c, y_c) - v_c)_y|$ is adjusted such that $v'_c = v_c + \text{sign}((F(x_c, y_c) - v_c)_y) \cdot y_{max}$ where y_{max} is the size of the image in the y -direction. The image is then shifted in the y -direction so as to minimise v_c and the full multi-level B-spline fitting is repeated to give the final function $F(\Phi)$ with control lattice Φ .

Now for every position in the prone image $P_{prone} = \{(x, y)\} \in \Omega$ we can use the function F to find the corresponding position in the supine image $P_{supine} = \{(x, y)\} \in \Omega$. We can use this transformation alone, or use it as an initialisation to the intensity-based B-spline registration function presented in [7] to create a finer composite registration.

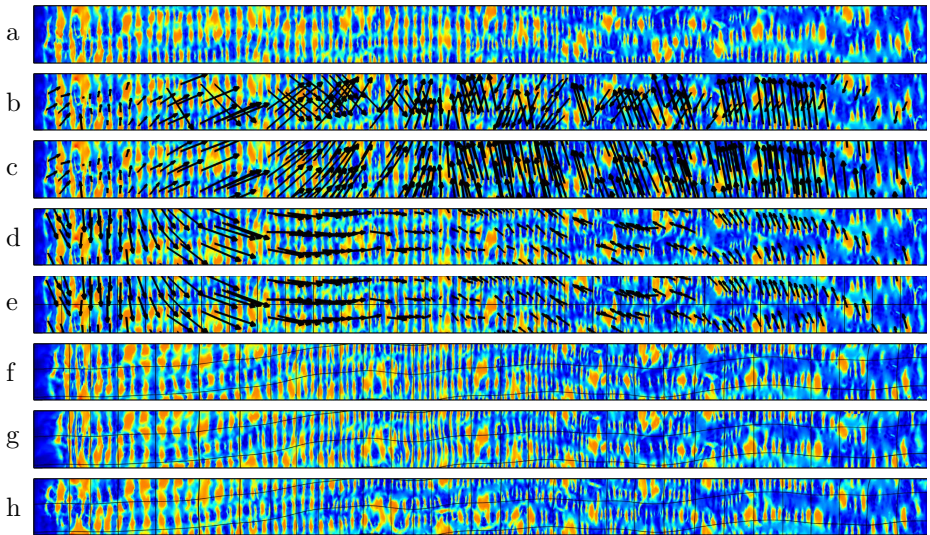


Fig. 1. Images of the endoluminal surface produced from the conformal mapping technique (case 11). The colour scheme shows the Shape Index (SI) and the vectors show the displacements generated from the landmark registration. Images show: a) the source (prone) image; b) the ambiguous vector direction on the source image; c) the sorted displacements; d) the source image vertically aligned to reduce displacements; e) the source image with displacement vectors and regular grid; f) the result of the landmark B-spline registration with transformed image and grid; g) the refinement with the intensity based registration (with same grid); h) the target (supine) image.

3 Evaluation

3.1 Data

Ethical approval and informed consent was obtained to use anonymised CT colonography data. Colonic cleansing and insufflations has been performed in accordance with current recommendations [11]. These data consist of 17 validation cases of which 5 exhibited local luminal collapse (see table 1), and 4 cases (cases 9 - 12) that had been excluded from a previous study using an intensity based registration [7] due to marked differences in local distension and therefore different surface features. Cases used fluid tagging (allowing for digital cleansing of residual fluid) or little fluid remained. All parameters were optimised on separate training data. A radiologist (experienced in over 400 validated colonography studies) manually identified the centres of corresponding haustral folds using 'virtual colonoscopy' fly-through renderings, and external views of the colonic lumen. This resulted in 1484 pairs of corresponding positions between the two views to be used for evaluation.

Table 1. Information of cases exhibiting local luminal collapse. For each case, the number of collapsed regions in the prone and supine images are displayed, along with the Euclidean distance across each region. Locations of collapse are given (DC: descending colon; SC: sigmoid colon).

Case No.	Prone			Supine		
	Collapses	Location	Distance (mm)	No. Collapses	Location	Distance (mm)
13	1	DC	65.0	0	-	-
14	1	DC	245.1	1	DC	272.4
15	0	-	-	1	SC	26.0
16	3	DC	6.5	0	-	-
		DC	34.4			
		SC	8.0			
17	0	-	-	1	DC	18.3

To assess the performance of each algorithm in terms of registration error, each reference standard point is transformed from prone to supine using the calculated transformation function and the 3D Euclidean distance between this and the corresponding reference standard point is measured. It is clear that the composite method outperforms both the intensity- and landmark-based registration methods at $7.0mm(\pm 2.7mm)$, compared to $12.1mm(\pm 10.6mm)$ and $9.2mm(\pm 3.4mm)$ respectively. Using a Related Samples Wilcoxon Signed Rank Test, the difference in error between each pair of results is statistically significant with $p < 0.001$.

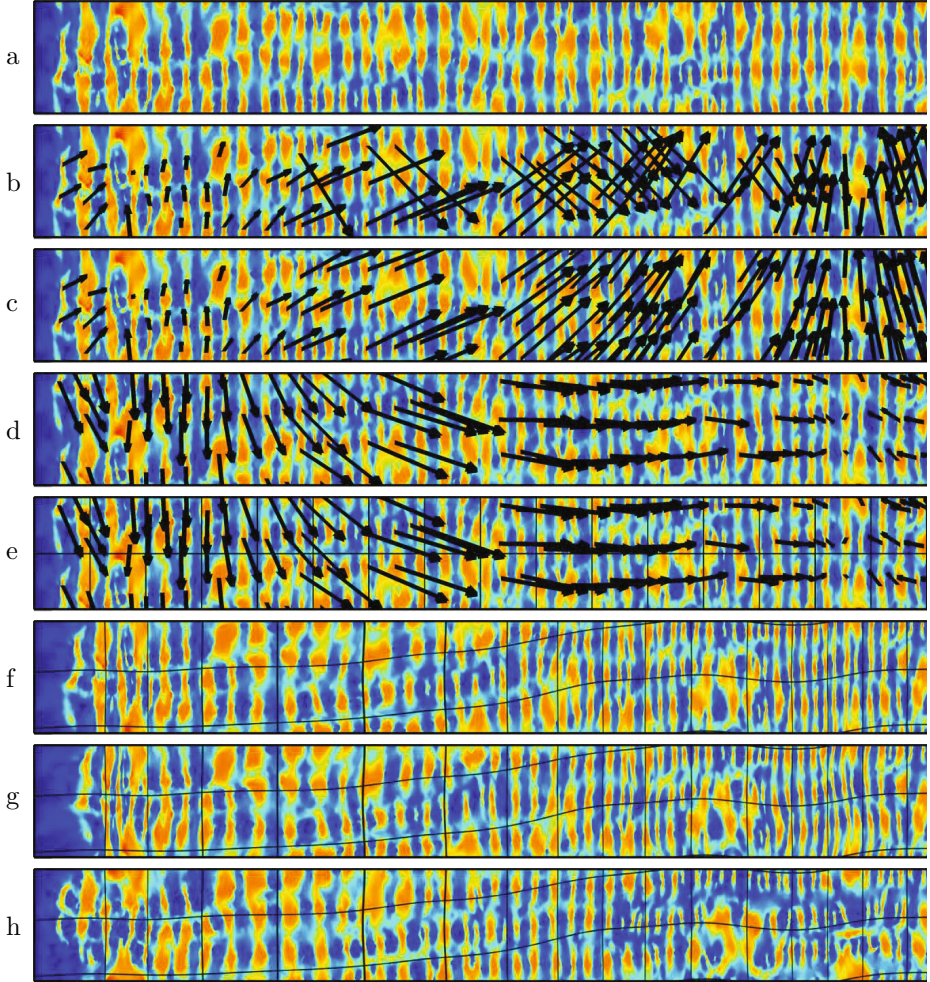


Fig. 2. Displaying the images in figure 1 in greater detail at the end of the colon. The colour scheme shows the Shape Index (SI) and the vectors show the displacements generated from the landmark registration. Images show: a) the source (prone) image, b) the ambiguous vector direction on the source image; c) the sorted displacements; d) the source image vertically aligned to reduce displacements; e) the source image with displacement vectors and regular grid; f) the result of the landmark B-spline registration with transformed image and grid; g) the refinement with the intensity based registration (with same grid); h) the target (supine) image.

Table 2. Mean fold registration error (mm) for each of the validation cases. Results are shown individually for the intensity, landmark and composite registration methods.

	Case	Intensity	Landmark	Composite	
Fully Distended	1	12.0	10.1	11.7	
	2	7.4	7.3	5.6	
	3	5.3	6.0	5.2	
	4	9.0	6.1	5.6	
	5	5.6	5.1	5.6	
	6	3.5	4.7	3.6	
	7	5.9	6.4	5.6	
	8	6.9	7.3	6.2	
	Subset Mean	7.0	6.6	6.2	
	Subset Std	2.6	1.7	2.4	
	Previously Excluded Cases				
	9	44.9	15.2	5.8	
10	12.5	8.5	7.8		
11	16.8	10.3	6.0		
12	7.3	10.3	6.2		
Subset Mean	20.4	11.1	6.5		
Subset Std	16.8	2.9	0.9		
Collapsed	13	30.3	16.2	15.0	
	14	5.9	13.6	6.8	
	15	7.5	11.1	7.2	
	16	15.7	9.1	6.7	
	17	8.8	8.4	7.6	
	Subset Mean	13.7	11.7	8.7	
	Subset Std	10.0	3.2	3.6	
Total Mean	12.1	9.2	7.0		
Total Std	10.6	3.4	2.7		

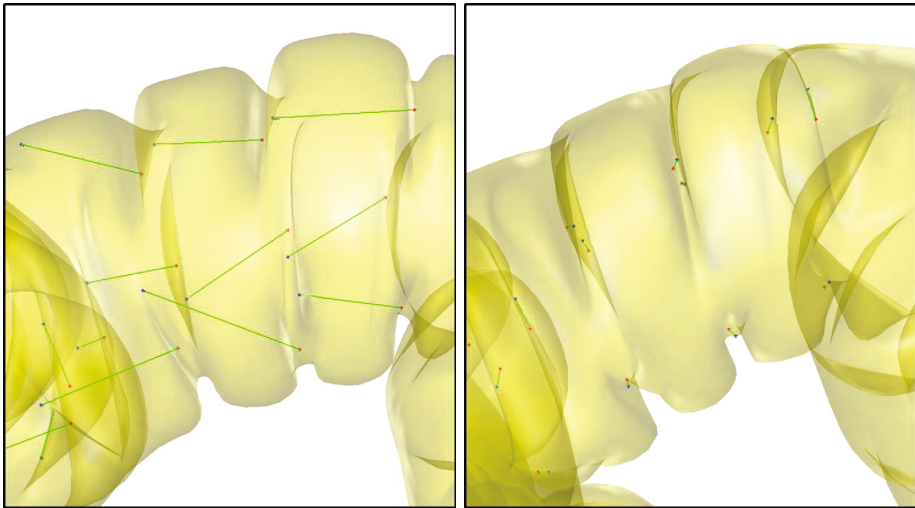


Fig. 3. External surface renderings of the transverse colon in the supine image of case 16. The set of reference standard points in the supine view (red) and the corresponding points transformed from the prone view (blue) and shown using the results from the intensity based (left) and composite (right) registration methods. The green lines show the Euclidean distance error.

4 Conclusion

Our composite registration method combines landmark and intensity based registration techniques and improves the mean registration accuracy compared to using either method alone. The work flow presented is fully automated, taking as input a prone and supine colon lumen segmentation. The consistency of results across cases showing a variety of characteristics indicates that the composite method will provide a more robust registration than those previously reported, especially in more 'difficult' cases, such as those that show marked differences in distension, or exhibit areas of local colonic collapse. This situation is very common in routine practice and therefore an algorithm that is robust to these characteristics is of greater clinical benefit.

References

1. Fukano, E., Oda, M., Kitasaka, T., Suenaga, Y., Takayama, T., Takabatake, H., Mori, M., Natori, H., Nawano, S., Mori, K.: Haustral fold registration in CT colonography and its application to registration of virtual stretched view of the colon. In: Proceedings of SPIE, vol. 7624, p. 762420 (2010)
2. Hampshire, T., Roth, H., Hu, M., Boone, D., Slabaugh, G., Punwani, S., Halligan, S., Hawkes, D.: Automatic Prone to Supine Haustral Fold Matching in CT Colonography Using a Markov Random Field Model. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 508–515. Springer, Heidelberg (2011)

3. Jin, M., Kim, J., Luo, F., Gu, X.: Discrete surface ricci flow. *IEEE Transactions on Visualization and Computer Graphics* 14(5), 1030–1043 (2008)
4. Koenderink, J.J.: *Solid shape*. MIT Press, Cambridge (1990)
5. Lee, S., Wolberg, G., Shin, S.Y.: Scattered data interpolation with multilevel B-splines. *IEEE Transactions on Visualization and Computer Graphics* 3(3), 228–244 (1997)
6. 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)
7. Roth, H., McClelland, J., Boone, B., Modat, M., Cardoso, M., Hampshire, T., Hu, M., Punwani, S., Ourselin, S., Slabaugh, G., Halligan, S., Hawkes, D.: Registration of the endoluminal surfaces of the colon derived from prone and supine CT colonography. *Medical Physics* 38(6), 3077–3089 (2011), <http://link.aip.org/link/?MPH/38/3077/1>
8. 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 Transactions on Medical Imaging* 18(8), 712–721 (1999)
9. Suh, J.W., Wyatt, C.L.: Deformable registration of supine and prone colons for computed tomographic colonography. *Journal of Computer Assisted Tomography* 33(6), 902 (2009)
10. Taylor, S.A., Halligan, S., Goh, V., Morley, S., Bassett, P., Atkin, W., Bartram, C.I.: Optimizing colonic distention for multi-detector row CT colonography: Effect of hyoscine butylbromide and rectal balloon catheter. *Radiology* 229(1), 99 (2003)
11. Taylor, S.A., Laghi, A., Lefere, P., Halligan, S., Stoker, J.: European society of gastrointestinal and abdominal radiology (esgar): consensus statement on CT colonography. *European Radiology* 17(2), 575–579 (2007)
12. Wang, S., Yao, J., Liu, J., Petrick, N., Van Uitert, R.L., Periaswamy, S., Summers, R.M.: Registration of prone and supine CT colonography scans using correlation optimized warping and canonical correlation analysis. *Medical Physics* 36, 5595 (2009)
13. Weiss, Y., Freeman, W.T.: On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs. *IEEE Transactions on Information Theory* 47(2), 736–744 (2002)
14. Zeng, W., Marino, J., Gurijala, K.C., Gu, X., Kaufman, A.: Supine and prone colon registration using quasi-conformal mapping. *IEEE Transactions on Visualization and Computer Graphics* 16, 1348–1357 (2010)
15. Zeng, W., Samaras, D., Gu, D.: Ricci flow for 3D shape analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(4), 662–677 (2010)

External Clinical Validation of Prone and Supine CT Colonography Registration

Holger R. Roth¹, Darren J. Boone², Steve Halligan², Thomas E. Hampshire¹,
Jamie R. McClelland¹, Mingxing Hu¹, Shonit Punwani²,
Stuart Taylor², and David J. Hawkes¹

¹ Centre for Medical Image Computing, University College London,
London, WC1E 6BT, UK
h.roth@ucl.ac.uk

² Centre for Medical Imaging, Department of Specialist Radiology,
University College Hospital, London, UK

Abstract. This paper provides an external validation of a prone-supine registration algorithm for CT colonography (CTC). A validation sample of 49 patient cases with 66 polyps (6 to 30 mm) was selected from a publicly available, anonymized CTC archive. To enhance generalizability, no case was excluded due to poor preparation or inadequate distension. Corresponding prone and supine polyp coordinates were recorded and the endoluminal surfaces registered: a Markov Random Field technique was used to find feature matches between prone/supine acquisitions and following mapping of the endoluminal surface to a cylinder, dense surface correspondence was achieved via cylindrical non-rigid registration. The polyp registration error was determined and a subjective assessment of registration made for 2D slice-based and 3D endoluminal data display using a pre-specified scoring system. Results were compared to using “normalized distance along the colon centerline” (NDACC) which approximates to the method currently employed to match colonic positions using proprietary CT colonography interpretation software. Registration was possible in all 49 cases. Overall mean 3D polyp registration error was significantly smaller with 19.9 mm in comparison to 27.7 mm using NDACC ($p=0.001$). 82.7% of polyp matches were defined as “successful” in comparison to 37.1% using NDACC according to the pre-specified criteria. Similarly, using 2D visualization, 62.1% registrations were “successful” and only 22.7% using NDACC. Full surface-based prone-to-supine registration can successfully map the location of a polyp identified on one acquisition to the corresponding endoluminal surface in the opposing acquisition, greatly facilitating polyp matching and aiding interpretation. Our method compares favorably to using NDACC.

Keywords: CT colonography, image registration.

1 Introduction

CT colonography interpretation is difficult and time consuming. Fecal residue can simulate or obscure polyps, leading to both false positive and false negative

diagnoses. To compensate for this, it is normal practice to obtain CT data with the patient both prone and supine to redistribute fecal residue and colonic gas; fecal residue tends to move, while fixed mural pathology does not. Matching corresponding endoluminal locations between prone and supine acquisitions is the cornerstone of competent interpretation [19]. Unfortunately, the colon often undergoes significant deformation during repositioning [12] which complicates the interpretative task and can induce reader error.

Software algorithms have attempted to facilitate matching of corresponding locations between prone and supine acquisitions: matching of distances along centerlines enables navigation to approximate endoluminal locations in both datasets [19,18,11,17]. The most straight-forward method (and the only technique available in commercial workstations at present) being the “normalized distance along the colon centerline” (NDACC) [16]. While centerline matching studies have shown promising results [19,18,11,17], they tend to report registration error in only one dimension. However, it can be argued that centerline-based methods are inherently limited because the interpretation task requires readers to find a specific point on the endoluminal surface rather than along the centerline. Moreover, many readers favor using multiplanar reformats rather than endoluminal 3D displays, further limiting the utility of centerline methods in routine daily practice.

Alternative methods that aim to register the full colonic surfaces have been proposed [15,14,6,21]. However, such studies often report the selection criteria incompletely [19,18,2,21,14] such that the transferability of results into clinical practice remains uncertain. For example, studies routinely use optimally prepared CT colonography datasets yet in reality, approximately 50% of cases are “poorly prepared” [5]. Likewise, validation should use data from centers that have not contributed to algorithm development (external validation) [3], to ensure previous exposure to the test data does not bias result.

We have developed software that establishes full spatial correspondence between the prone and supine endoluminal surfaces. We aim to externally validate this registration algorithm using CT colonography data generalizable to clinical practice and compare the results to those obtained using NDACC.

2 Prone to Supine Registration

The registration method has been described in detail previously [13]. Initialization is provided by robust matches of haustral folds between both views [4]. The haustral folds are segmented on the colon surface using a graph cut method applied to a surface curvature-based metric. Using a Markov Random Field based on similarity of folds and additional neighborhood information, fold matches are then estimated between both views. Only inverse-consistent matches were used, i.e. the same match is obtained supine-to-prone and prone-to-supine. Secondly, the entire endoluminal surface is mapped to a cylinder utilizing a conformal mapping based on Ricci flow [7]. The original surface curvature information is preserved during this step. Surface correspondence is then achieved using a non-rigid, cylindrical version of the well-known B-spline registration [10]. Registration

is driven by local shape measurements, i.e. shape index (SI) computed on the colon surface. This measure has also shown to be useful in computer-aided detection (CAD) for CT colonography [20]. The sum-of-squared differences (SSD) of these SI measures are used to drive the cylindrical registration. In case of insufficient colonic distension, any locally collapsed region can be ignored during this cylindrical registration when computing the similarity measure. This is important as collapse can commonly occur in clinical practice. We arbitrarily denote the prone image as target and supine image as source when computing the registration. After convergence of the algorithm, any point on the endoluminal surface can be mapped between both corresponding acquisitions in either direction (prone-to-supine and supine-to-prone) by inverting the result [13,4]. Fig. 1 illustrates the principle of this registration method.

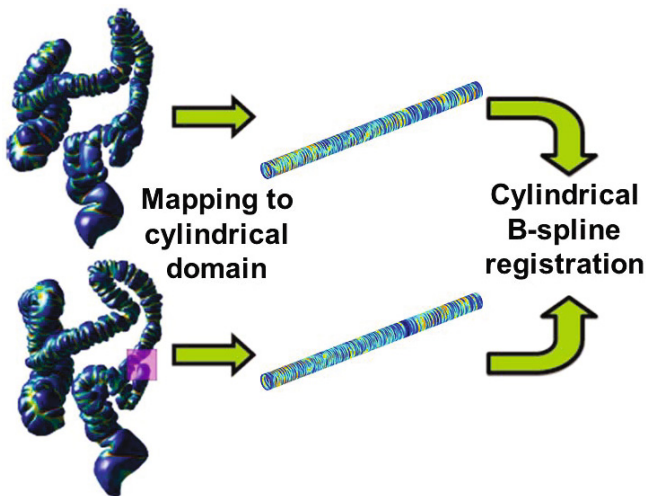


Fig. 1. Prone (top) to supine (bottom) registration of endoluminal colon surfaces in cylindrical space. The color coding indicates the local shape index SI measurements which drive a non-rigid cylindrical B-spline registration – even in the case of local collapse (indicated by square).

3 Methods and Materials

The external validation is based on publicly available CTC data [8]; the ACIN CT trial [1] investigators have made 103 anonymized cases, from 15 US centers, with polyps (including reference standard) available on the National Cancer Institutes National Biomedical Imaging Archive (NBIA) [2]. These cases include 69 medium (6-9 mm) and 32 large (≥ 10 mm) polyps. These data represents

¹ <http://www.acrin.org/TabID/151/Default.aspx>

² <https://imaging.nci.nih.gov/ncia/>

cases that are generalizable to daily practice and were uninvolved with the development of the registration method. No attempt has been performed to select ‘perfect’ cases; cases were examined and excluded if: no independent CTC reference data were available (41), the polyp was not visible in both datasets (7), the dataset was incomplete/corrupted (3) or the cancer/polyp was $> 30\text{ mm}$ (3). This leaves 49 cases with 66 polyps (38 large, 28 medium) used for validation. Each present polyp was identified by an experienced observer (DJB) using the independent reference data and prone and supine 3D polyp volumes were labeled manually using *ITK-snap*³. The segmental distribution of polyps in the validation sample (n=66) were compared to polyps $\geq 6\text{ mm}$ (n=547) from the entire ACRIN CTC study (21) (n=2525) to investigate the likely generalizability of our results. By adopting the criteria proposed by Hara et al. [5], 55% of validation cases (n=27) had excess residual fluid compared to 52% (1313) of the total CT colonography studies from the same trial. 47% (23) had at least one region of complete luminal collapse, similar to the 48% (50) observed in the total, 103, positive cases in the publicly available database.

4 Assessment of Clinical Utility

Currently, standard clinical workstations for CT colonography interpretation allow the rendering of a *virtual endoscopic* or *virtual fly-through* views in order to replicate the view during colonoscopy. We can judge the usefulness of our proposed registration algorithm in clinical practice using these three-dimensional renderings. To score the registration result, a polyp conspicuity grading was developed by two experienced radiologists (DJB, SH): after successfully establishing the prone-to-supine correspondence over the full colon surface, the user is automatically guided to the corresponding view in opposing data set when clicking on a potential polyp (see Fig. 2). The endoluminal display is generated using a standard 120° field-of-view camera model and centering the camera origin at the corresponding centerline position, pointing towards the surface registration result. The clinical usefulness is then judged by using the following objective polyp conspicuity score. The score is proportional to the amount of ‘mouse-driven navigation’ necessary in order to bring the target polyp into view after registration:

- **5 ‘Successful’**: Polyp directly marked by registration prompt (Fig. 2)
- **4 ‘Successful’**: Polyp within same field of view (Fig. 3)
- **3 ‘Partially successful’**: Polyp becomes visible after $\pm 90^\circ$ pan
- **2 ‘Partially successful’**: Polyp becomes visible after 360° pan
- **1 ‘Unsuccessful’**: Polyp only comes into view after navigation along centerline

Similarly, a 3D conspicuity assessment is performed in order to compare to NDACC. Here, we combine the scores of category 4 and 5 as NDACC will not

³ www.itksnap.org

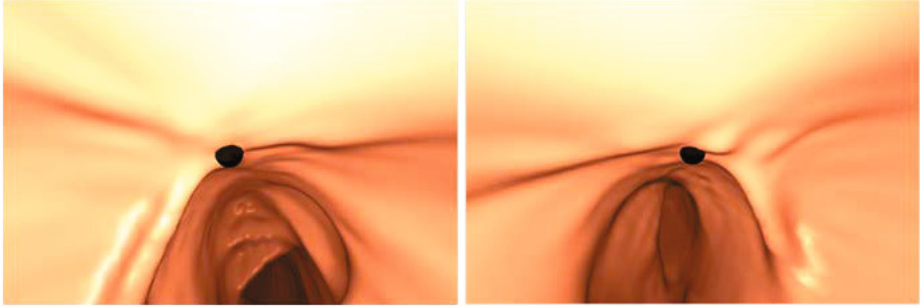


Fig. 2. Example of polyp conspicuity score of 5 ('direct hit'). The registration prompt (black dot) marks the polyp indicated by the observer in the supine dataset (left). Following registration, the algorithm centers the prone view to point towards the endoluminal coordinates calculated by the algorithm (right).

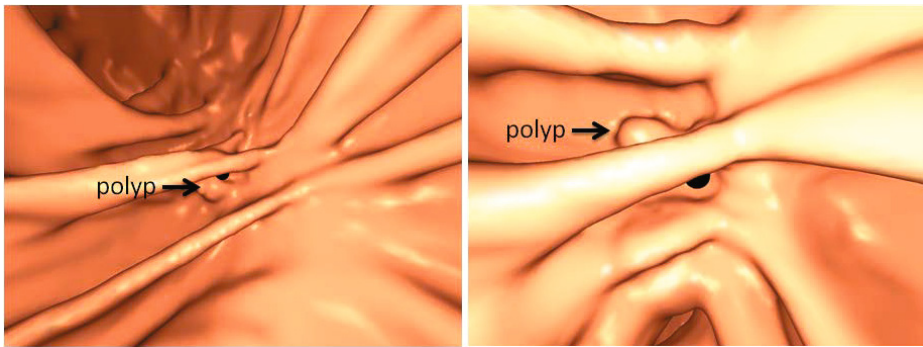


Fig. 3. Example of polyp conspicuity score of 4 ('near miss'). The registration prompt (black dot) marks the polyp indicated by the observer in the supine dataset (left). Following registration, the algorithm centers the prone 3D field of view to point towards the coordinates calculated by the algorithm (right). The algorithm fails to indicate the polyp (arrow) correctly but the polyp is clearly visible in the same field of view.

provide a registration prompt on the colon surface. It will only synchronize the virtual endoscopic views, looking straight along the centerline.

Furthermore, a conspicuity score was developed for aligning polyps using a standard multiplanar slice viewer for increasing necessity of mouse-driven navigation in all three orthogonal directions (axial, sagittal, coronal):

- 5 **'Successful'**: Polyp directly marked by registration prompt.
- 4 **'Successful'**: Polyp visible after 15 mm mouse-navigation
- 3 **'Partially successful'**: Polyp visible after ± 20 mm mouse-navigation
- 2 **'Partially successful'**: Polyp visible within ± 30 mm mouse-navigation
- 1 **'Unsuccessful'**: Polyp not within ± 30 mm of registration prompt

The polyp's apex was computed using the manual segmentations and was used to compute the 3D registration error and 2D slice displacement (measured in all 3 orthogonal directions). Both errors are reported in millimeters (mm).

Polyps could be in the field of view or not depending on the particular local configuration of the endoluminal colon shape and centerline. In order to reduce any bias towards the direction of transformation, all polyps were transformed in prone-to-supine as well as in supine-to-prone directions. This resulted in a total of 132 point transformations of 66 polyps for the following analysis. A 5% significance level was applied to all following cases when computing statistical significance. Despite the data reflecting clinical practice and depicting many areas of local colonic collapse, the registration was successfully computed in all of the cases (100%).

5 Results

3D Registration Error. The mean of the 3D registration error was 19.9 mm over the whole data set (66 polyps in 49 cases). This error measures the distance between the target polyp's apex and the point indicated on the colon surface by our proposed method. A median error of 11.9 mm was achieved and the 3D registration error ranged in the limits of 1.0 mm to 85.8 mm. The 3D registration error did not vary significantly when comparing the polyp's location in different colonic segments ($p=0.65$). Furthermore, there was no significant increase in registration error within data cases exhibiting at least one local luminal collapse ($p=0.075$) using a Kruskal-Wallis test.

In order to provide a comparison to how a NDACC method would be used in clinical practice, the polyp's position along the centerline is found as the minimum distance between the polyps position on the colonic wall and all points along the centerline, following the approach by Wang et al. [18]. In comparison, using NDACC leads to a mean 3D error of 27.7 mm between apex of the polyp and the point indicated along the centerline. Here, our measured error is significantly smaller ($p=0.001$).

1D Registration Error. Finding the polyp's position along the centerline further allows measuring the registration error as a 1D distance along centerline, summarized in Table 1. Using a non-parametric paired analysis (Wilcoxon Signed-Rank Test), our method gives no significant advantage over using NDACC when using all 66 polyps: 17.6 mm compared to 20.8 mm using NDACC. However, when comparing registration by colonic segment, the method shows significant improvement over NDACC in the transverse, descending and sigmoid colon (see Table 1). A mean of 19.3 mm is achieved for these colonic segments in comparison 26.9 mm (NDACC). This is a significant improvement with $p=0.047$. The largest improvement occurred in the transverse colon with a 15.3 mm reduction in registration error. This was anticipated as this segment was reported as being most mobile during positional change of the patient [12].

Table 1. 1D registration error, measured as displacement in along centerline [mm]

Colonic segment	Number of polyps	1D error proposed method	1D error NDACC
Rectum	14	13.1	14.7
Sigmoid	15	20.9	23.6
Descending	11	18.1	27.5
Transverse	7	17.9	33.2
Ascending	12	22.7	17.2
Cecum	7	9.9	10.5
Total	66	17.6	20.8

Polyp Conspicuity Score in 3D. The registration achieved an direct alignment of the registration prompt in both opposing data sets in 89 (67.4%) of cases using virtual fly-throughs (conspicuity score of 5; see Fig. 2). Of 132 point transformations, 20 (15.2%) were transformed within the same field of view, leading to a score of 4 (see Fig. 3). Using the conspicuity grading, this leads to 82.7% ‘successful’ registrations. Nine (6.8%) polyp transformations required up to 90 degrees mouse-driven rotation around the camera center in order to bring the polyp into view. In three accounts (2.3%), a full 360 degrees of rotation were required to achieve the same. Therefore, a total of 9.1% were rated being aligned ‘partially successful’ for clinical relevance. A total of eleven (8.4%) point transformations failed as further navigation along the centerline was necessary in order to bring the polyp into view. This compares favorable to using NDACC where a ‘successful’ (score 4) was only achieved in 37.1% of point transformations.

Polyp Conspicuity Score in 2D. Using a purely multiplanar 2D slice 2D visualization, the registration error was assessed by judging utility expressed in necessary mouse-driven navigation in order to align the polyp apices in all three orthogonal directions. The algorithm correctly aligned the slices of 82 (62.1%) pairs of prone and supine polyps within ± 15 mm of displacement and were therefore being judged ‘successful’ for 2D navigation. 28 (21.3%) of cases were ‘partially successful’ registered, meaning that they were aligned within ± 30 mm. A total of 22 registrations (16.7%) failed as more than ± 30 mm of navigation was necessary in order to bring them into view. In comparison, NDACC achieved only 30 (22.7%) ‘successful’ alignments within ± 15 mm of displacement. 78 (59.1%) were judged ‘partially successful’ and 24 (18.2%) failed and was not visible within ± 30 mm of navigation.

6 Discussion

In current clinical practice, prior to virtual colonoscopy, the interpreting clinician must check the validity of the workstation’s proposed colonic segmentation, primarily to ensure the computed centerline takes the correct path to ensure

correct navigation along the colonic lumen. Likewise, our registration algorithm currently relies on the user accepting the colon segmentation and the correct order of colonic segments, defining start and end points. Thereafter, the algorithm for establishing correspondence over the whole colonic surface is automated.

The assessment of registration accuracy used data from several hospitals unrelated to the development of the algorithm. This enables a generalizable assessment of the algorithm's performance which could be applied to normal clinical practice. Our study sample closely parallels the ACRIN data in quality of the bowel preparation and distension. Our algorithm compares favorably with previous methods which use optimally prepared cases or exclude cases exhibiting local collapses for validation. We provide full surface registration of the colonographic data that enables the indication of a specific location on the endoluminal surface. This would provide the observer with considerably more information, compared to a centerline position from where further navigation is necessary in order to find the target location. This is reflected in our choice of pre-specified criteria for polyp conspicuity using 3D endoscopic fly-through visualization following registration. Measuring the amount of navigation necessary in order to align prone and supine polyps after registration allows a direct comparison to centerline-based matching using NDACC. While it is intuitive that an accurate prone to supine registration would be beneficial for shortening interpretation time, further studies of clinical utility in everyday practice are necessary. For example, the effect on sensitivity and specificity when finding polyps needs to be evaluated.

This study is not without limitations: Some studies from the ACRIN validation archive were necessarily excluded due to inability of our radiologists to confirm the presence of a polyp in both the prone and supine acquisitions. It is therefore possible that some, inadequately distended studies were excluded from the analysis. Nevertheless, we postulate that the NDACC methods, due to its reliance on a continuous colonic segmentation, without regions of collapse, would have fared even worse than the proposed method when dealing with these data. In addition, while the processing of the endoluminal centerline occurs rapidly and is available to the reporting radiologist at the time of reporting, our method, at present, requires considerable processing time for surface extraction, feature matching and non-rigid registration. However, once this step has been complete, transferring between matched prone and supine locations is practically instantaneous. This is reminiscent of another, well-established adjunct to CT colonographic interpretation – computer-aided detection (CAD) which, although now integrated into many vendor platforms, initially required several hours of processing time. We anticipate that with the integration of multi-threading and/or GPU-technology we can reduce this such that integration into routine work-flow could be possible.

In summary, we successfully applied computer-assisted prone-supine registration of specific points (polyps) on the endoluminal surface. The used CT colonography data used for validation was from a subset of a larger multi-center trial. Resulting registration accuracies show promise of the ability to rapidly and

automatically match switch between prone and supine positions and compare favorably to using NDACC. This will further facilitate the interpretation of CT colonography data.

Acknowledgements. The authors gratefully acknowledge financial support for this work from the NIHR program: *Imaging diagnosis of colorectal cancer: Interventions for efficient and acceptable diagnosis in symptomatic and screening populations* (Grant No. RP-PG-0407-10338) and the *EPSRC-CRUK Comprehensive Cancer Imaging Centre of UCL and KCL* (Grant No. C1519AO). Image data used in this research were obtained from *The Cancer Imaging Archive*⁴ sponsored by the *Cancer Imaging Program, DCTD/NCI/NIH*.

References

1. Acar, B., Napel, S., Paik, D., Li, P., Yee, J., Beaulieu, C., Jeffrey, R.: Registration of supine and prone CT colonography data: Method and evaluation. *Radiology* 221, 332 (2001)
2. Fukano, E., Oda, M., Kitasaka, T., Suenaga, Y., Takayama, T., Takabatake, H., Mori, M., Natori, H., Nawano, S., Mori, K.: Hausstral fold registration in CT colonography and its application to registration of virtual stretched view of the colon. In: *Proc. SPIE Med. Imaging*, vol. 7624, p. 762420 (2010)
3. Halligan, S., Taylor, S., Dehmeshki, J., Amin, H., Ye, X., Tsang, J., Roddie, M.: Computer-assisted detection for CT colonography: external validation. *Clin. Radiol.* 61, 758–763 (2006)
4. Hampshire, T., Roth, H., Hu, M., Boone, D., Slabaugh, G., Punwani, S., Halligan, S., Hawkes, D.: Automatic Prone to Supine Hausstral Fold Matching in CT Colonography Using a Markov Random Field Model. In: Fichtinger, G., Martel, A., Peters, T. (eds.) *MICCAI 2011, Part I. LNCS*, vol. 6891, pp. 508–515. Springer, Heidelberg (2011)
5. Hara, A., Kuo, M., Blevins, M., Chen, M., Yee, J., Dachman, A., Menias, C., Siewert, B., Cheema, J., Obregon, R., et al.: National CT colonography trial (ACRIN 6664): comparison of three full-laxative bowel preparations in more than 2500 average-risk patients. *Am. J. Roentgenol.* 196, 1076–1082 (2011)
6. Huang, A., Roy, D., Summers, R., Franaszek, M., Petrick, N., Choi, J., Pickhardt, P.: Teniae coli-based circumferential localization system for CT colonography: Feasibility study. *Radiology* 243, 551–560 (2007)
7. Jin, M., Kim, J., Luo, F., Gu, X.: Discrete surface ricci flow. *IEEE Trans. Vis. Comput. Graphics* 14, 1030–1043 (2008)
8. Johnson, C., Chen, M., Toledano, A., Heiken, J., Dachman, A., Kuo, M., Menias, C., Siewert, B., Cheema, J., Obregon, R., et al.: Accuracy of CT colonography for detection of large adenomas and cancers. *N. Engl. J. Med.* 359, 1207–1217 (2008)
9. Li, P., Napel, S., Acar, B., Paik, D., Jeffrey Jr., R., Beaulieu, C.: Registration of central paths and colonic polyps between supine and prone scans in computed tomography colonography: Pilot study. *Med. Phys.* 31, 2912–2923 (2004)

⁴ <http://cancerimagingarchive.net/>

10. Modat, M., Ridgway, G., Taylor, Z., Lehmann, M., Barnes, J., Hawkes, D., Fox, N., Ourselin, S.: Fast free-form deformation using graphics processing units. *Comput. Meth. Prog. Bio.* 98, 278–284 (2010)
11. Näppi, J., Okamura, A., Frimmel, H., Dachman, A., Yoshida, H.: Region-based supine-prone correspondence for reduction of false-positive CAD polyp candidates in CT colonography. *Acad. Radiol.* 12, 695–707 (2005)
12. Punwani, S., Halligan, S., Tolan, D., Taylor, S., Hawkes, D.: Quantitative assessment of colonic movement between prone and supine patient positions during CT colonography. *Br. J. Radiol.* 82, 475–481 (2009)
13. Roth, H., McClelland, J., Boone, D., Modat, M., Cardoso, M., Hampshire, T., Hu, M., Punwani, S., Ourselin, S., Slabaugh, G., Halligan, S., Hawkes, D.: Registration of the endoluminal surfaces of the colon derived from prone and supine CT colonography. *Med. Phys.* 38, 3077–3089 (2011)
14. Suh, J., Wyatt, C.: Deformable registration of supine and prone colons for computed tomographic colonography. *J. Comput. Assist. Tomogr.* 33, 902–911 (2009)
15. Suh, J., Wyatt, C.: Registration of prone and supine colons in the presence of topological changes. In: *Proc. SPIE Med. Imaging*, vol. 6916, p. 69160C (2008)
16. Summers, R., Swift, J., Dwyer, A., Choi, J., Pickhardt, P.: Normalized distance along the colon centerline: a method for correlating polyp location on CT colonography and optical colonoscopy. *Am. J. Roentgenol.* 193, 1296–1304 (2009)
17. de Vries, A., Truyen, R., Van der Peijl, J., Florie, J., van Gelder, R., Gerritsen, F., Stoker, J.: Feasibility of automated matching of supine and prone CT-colonography examinations. *Br. J. Radiol.* 79, 740–744 (2006)
18. Wang, S., Yao, J., Liu, J., Petrick, N., Van Uitert, R., Periaswamy, S., Summers, R.: Registration of prone and supine CT colonography scans using correlation optimized warping and canonical correlation analysis. *Med. Phys.* 36, 5595–5603 (2009)
19. Yee, J., Kumar, N., Hung, R., Akerkar, G., Kumar, P., Wall, S.: Comparison of supine and prone scanning separately and in combination at CT colonography. *Radiology* 226, 653–661 (2003)
20. Yoshida, H., Näppi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *IEEE Trans. Med. Imaging* 20, 1261–1274 (2001)
21. Zeng, W., Marino, J., Chaitanya Gurijala, K., Gu, X., Kaufman, A.: Supine and prone colon registration using quasi-conformal mapping. *IEEE Trans. Vis. Comput. Graph* 16, 1348–1357 (2010)

Efficient Topological Cleaning for Visual Colon Surface Flattening

Rui Shi¹, Wei Zeng², Jerome Zhengrong Liang¹, and Xianfeng David Gu¹

¹ Department of Computer Science, Department of Radiology,
Stony Brook University, Stony Brook, NY 11794, USA
rshi@cs.sunysb.edu

² School of Computing & Information Sciences, Florida International University,
11200 SW 71st, Miami, FL 33199, USA

Abstract. Conformal mapping provides a unique way to flatten the three dimensional (3D) anatomically-complicated colon wall. Visualizing the flattened 2D colon wall supplies an alternative means for the task of detecting abnormality as compared to the conventional endoscopic views. In addition to the visualization, the flattened colon wall carries supplementary geometry and texture information for computer aided detection of abnormality. It is hypothesized that utilizing both the original 3D and the flattened 2D colon walls shall improve the detection capacity of currently available computed tomography colonography. One of the major challenges for the conformal colon flattening is how to make the input colon wall inner surface to be genus zero, as this is required by the flatten algorithm and will guarantee high flatten quality. This paper describes an efficient topological cleaning algorithm for the conformal colon flattening pipeline. Starting from a segmented colon wall, the Marching Cube algorithm was first applied to generate the surface, then we apply our topological clearance algorithm to remove the topological outliers to guarantee the output surface is exactly genus 0. The cleared or denoised colon surface was then flattened by an Ricci flow. The pipeline was tested by 14 patient datasets with comparison to our previous work.

Keywords: Flattening, conformal mapping, homotopy, Ricci flow, virtual colonoscopy.

1 Introduction

Virtual colonoscopy (VC), mimicking the conventional optical colonoscopy (OC), is a medical imaging procedure which uses X-rays or magnetic resonance (MR) signals and computers (1) to produce two and three-dimensional (3D) images of the colon (large intestine) from the lowest part, i.e., the rectum, all the way toward the lower end (i.e., the cecum) of the small intestine, and (2) to visualize the colon mucosal surface by endoscopic views on a screen [7,8]. The procedure has shown the potential to screen colonic polyps and detect colon diseases, including diverticulosis and cancer [14].

Traditional paradigm in VC employs X-ray computed tomography or computed tomography colonography (CTC) to achieve the tasks of screening and detection due to the high speed of CT scanning and high contrast between colon wall and colon lumen filled by CO₂ or air in CT images. While MR colonography (MRC) has an attractive

point of non-ionization radiation [16], it faces several drawbacks, e.g., lower spatial resolution, prone to motion artifacts, and noticeable susceptibility artifacts on the interface between air and tissue/colon wall. Therefore, MRC remains in the early research development stage, while CTC has been successfully demonstrated to be more convenient and efficient than OC as a screening modality [14]. A combination of VC screening with OC follow-up for therapeutic intervention could be a cost-effective means to prevent the deadly disease of colon cancer.

However, because of the length of the colon with complicated structures, inspecting the entire colon wall is time consuming and prone to errors by current VC technologies. Moreover, because of the complicated colon structure, the field-of-view (FOV) of the VC endoscopic views is limited, resulting in incomplete examinations. Flattening the 3D wall into a 2D image would effectively increase the FOV and provide supplementary information to the VC endoscopic views [5]. Thereafter, various flattening techniques [2,3,9,13,10,18] have been developed, among which the conformal mapping algorithm [9,10] showed advantages in generating 2D colon wall image with minimal distortion by preserving the structural angles.

Paik et al. [13] used cartographic projections to project the whole solid angle of the camera. This approach samples the solid angle of the camera, and maps it onto a cylinder which is finally mapped to a 2D planar image. However, this method causes distortions in shape. Bartrol et al. [3] moved a camera along the central path of the colon. However, this approach does not provide a complete overview of the colon. Haker et al. [5] employed certain angle-preserving mappings, based on a discretization of the Laplace-Beltrami operator, to flatten a surface onto a plane in a manner which preserves the local geometry. However, the flattened result of their method is not efficient for applications like polyp identification, and it requires a highly accurate and smooth surface mesh to achieve a good mean-curvature calculation. Wang, et al. [18] explored a volume-based flattening strategy to visualize the textures of the original 3D colon wall in the flattened 2D image. However, the distance-based mapping may not be accurate enough for detection of small polyps. The associated computation is intensive.

Hong et al. [9,15] utilized conformal structure to flatten the colon wall onto a planar image. Their method is angle preserving and minimizes the global distortion. First, the colon wall is segmented and extracted from the CTC image data set. The topology noise (i.e., minute handle) is removed by a volumetric algorithm. The holomorphic 1-form, a pair of orthogonal vector fields, is then computed on the 3D colon surface mesh using the conjugate gradient method. The colon surface is cut along a vertical trajectory traced using the holomorphic 1-form. Consequently, the 3D colon surface is conformal mapped onto a 2D rectangle. The flattened 2D mesh is then rendered using a direct volume rendering method accelerated with the GPU strategy. For applications like polyp detection, the shape of the polyps is well preserved on the flattened colon images, and thus provides an efficient way to enhance the navigation of a virtual colonoscopy system.

Unfortunately, the de-noise algorithm in [9,15] cannot always get genus 0 surface (actually only one case succeed out of 14). In this paper, topological de-noise is solved by our new algorithm, which guarantees the output surface to be genus 0. This efficient de-noise algorithm greatly improved the efficiency and accuracy and deliver comparable flattening results.

2 Method

Fig. 1 shows our new conformal mapping pipeline with comparison to our previous one. From acquired CTC datasets, our first task was to segment each image data volume and extract the corresponding colon wall. This was achieved by a statistical maximum a posteriori expectation-maximization (MAP-EM) algorithm [17]. Both our present and previous pipelines share this task. Then a triangle mesh of the colon wall mucosal surface was generated by the standard Marching Cube algorithm. To remove topology handles, a new surface-based de-noise algorithm was applied. In our present pipeline, the Marching Cube algorithm was applied prior to topological de-noising. After de-noising, we developed Ricci flow method to perform the conformal flattening task.

Conformal mapping has many unique properties in flattening the colon wall, as shown in [9]. However, as we mentioned, the old de-noise method cannot guarantee the output surface is genus 0. Our contribution: we developed and applied a new topological de-noise algorithm, which is very efficient and can guarantee the output surface to be genus 0.

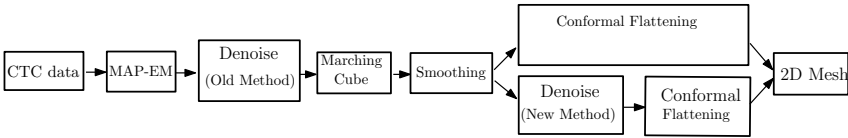


Fig. 1. Pipeline for our previous and current methods

3 Theoretic Background

This section briefly introduces elementary theories of surface topology and surface Ricci Flow.

3.1 Homotopy Basis

Definition 1 (Homotopy). Two continuous maps $f_0, f_1 : M \rightarrow N$ are said to be homotopic if there is a continuous map $f : M \times I \rightarrow N$ such that $F(\cdot, 0) = f_0$ and $F(\cdot, 1) = f_1$. The map F is called a homotopy [12] between f_0 and f_1 , denoted as $f_0 \cong f_1$.

Definition 2 (Homotopic Paths) [1]. Two paths f, g in M are said to be equivalent if f and g are homotopic relative to $\{0, 1\}$, denoted as $f \cong g$.

Definition 3 (Homotopy Basis). For a closed orientable surface M with genus g (i.e., a torus with g handles), there are $2g$ classes of homologically independent loops, called the homotopy basis of the surface M . A homology basis consists of one loop from each class, as shown in figure 2.

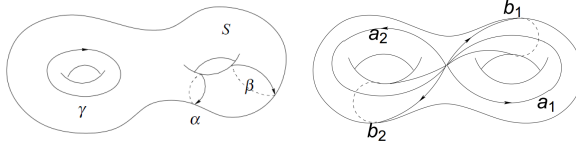


Fig. 2. Left: Homotopy: Curve α is homotopy to β , but not homotopy to γ . Right: Homotopy basis for a genus 2 surface. The number of loops $L = 2g = 4 \{a_1, a_2, b_1, b_2\}$, each as a representative of it's homotopy class.

3.2 Surface Ricci Flow

Suppose S is a surface in three dimensional Euclidean space \mathbb{R}^3 , therefore it has naturally the induced Euclidean metric \mathbf{g} . The Gaussian curvature is determined by the Riemannian metric \mathbf{g} , and satisfies the following Gauss-Bonnet theorem:

Theorem 1 (Gauss-Bonnet Theorem). *The total Gaussian curvature of a closed metric surface is*

$$\int_S K dA = 2\pi\chi(S),$$

where $\chi(S)$ is the Euler number, which equals to $\chi(S) = 2 - 2g$ for closed surface with genus g .

Ricci flow is a powerful curvature flow method invented by Hamilton[6] for the purpose of proving Poincaré's conjecture. Intuitively, it describes the process to deform the Riemannian metric according to curvature such that the curvature evolves like a heat diffusion process:

$$\frac{d\mathbf{g}}{dt} = -2K\mathbf{g}. \quad (1)$$

The convergence of surface Ricci flow was also proved in[6].

Theorem 2. *Suppose S is a closed surface with a Riemannian metric. If the total area is preserved, the surface Ricci flow will converge to a Riemannian metric of constant Gaussian curvature[6].*

Fig. 3 shows the conformal parameterization of colon surface computed by Ricci flow, figure 4 shows that conformal mapping preserves angle.

4 Algorithm

4.1 Topological De-noise

In previous work [9], the de-noise process started from the segmentation of the colon, incorporated the simple point concept in a region growing based algorithm to extract a topologically simple segmentation of the colon lumen. However, as we mentioned, the de-noise algorithm cannot guarantee the output surface is genus 0 in practice, so we developed a new efficient surface based de-noise algorithm to remove the topological

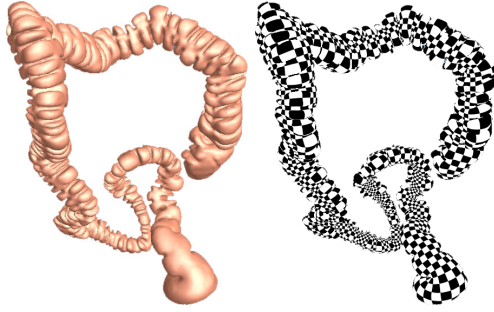


Fig. 3. Conformal parameterization of the colon surface

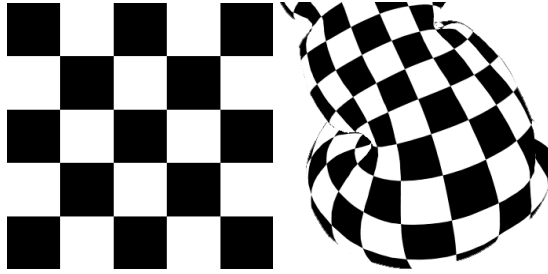


Fig. 4. Conformal Mapping preserves angle: the angle on the texture domain is well preserved after mapped to the surface, so the shape information of colon surface is well preserved

Algorithm 1. Topological De-noise Algorithm

Input: Surfaces M .

Output: Genus 0 surface \bar{M} .

1. Compute the homotopy basis G of M using algorithm 2.
 2. For each point p on homotopy basis G , grow a patch P .
 3. Find the shortest homotopy loop p_l starts at p in patch P .
 4. Find the shortest loop $\min\{p_l\}$ among all the vertices on G .
 5. Cut M along $\min\{p_l\}$ and fill the 2 holes appeared, get \bar{M} .
 6. If \bar{M} is not genus 0, goto step 1.
-

noise. As our method find tiny handles based on surface topology and remove them one by one, the final surface is guaranteed to be genus 0. The pipeline is like the following:

The idea of efficient topological de-noise algorithm is like following: we can compute the shortest loop goes though vertex v for all the vertices in mesh M , then find the shortest one among them, it must be the shortest handle loop in M . Furthermore, all the handles of a surface must be “go around” by the homotopy basis. As a result, we just need to compute the shortest loops for vertices on the homotopy basis G instead of all vertices of surface M , which leads to a 10 times speed up. Compared to the old voxel

based de-noise algorithm, our surface based method is much faster, and guarantees the output surface to be exactly genus 0. Fig. 5 shows a tiny handle went though by the homotopy basis.

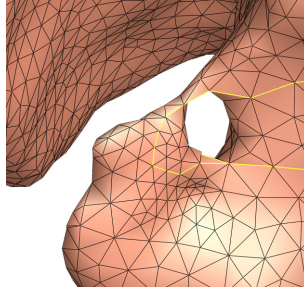


Fig. 5. A typical tiny hole and the homotopy basis (labeled by yellow line) goes around it

A homotopy basis at s can be thought of as a homology basis where all loops meet at a common vertex s , called the basepoint. Erickson and Whittlesey [4] proved that a shortest homotopy basis at a point on a mesh with n vertices can be computed in $O(n \log n)$ time. Fig. 6 shows the homotopy basis for genus 1 surface, algorithm 2 gives the algorithm for computing the homotopy basis:

Algorithm 2. Homotopy Basis Algorithm

Input: Surfaces M .

Output: The homotopy basis G of \bar{M} .

1. Find the maximum spanning tree T from a basepoint s .
 2. Find a maximum spanning tree T^* on the edges of the dual graph which do not cross edges in T .
 3. Find all edges $\{e_1, e_2 \dots e_{2g}\}$ which are neither in T nor are crossed by edges in T^* .
 4. Find the loops containing each e_i (using T), these loops form the homotopy basis.
-

4.2 Discrete Ricci Flow

The computation of the conformal mapping of a triangular mesh is based on the discrete Ricci flow [11][9], as Algorithm 3 shows.

5 Experimental Results

CTC datasets was random selected from a database. The presented algorithm was implemented in a similar manner as the previous algorithm [9] for a fair comparison. These algorithms were executed on a PC platform of Intel Xeon X5450 3.0GHz CPU and 8.00 GB RAM. To get the maximum quality, we use the original un-simplified mesh for conformal mapping. The triangle number of the 14 datasets range from 700k to 1200k. The method in [9] can only find around half of the handles, while our method can completely remove all the handles. Table 1 shows the de-noise result comparison between

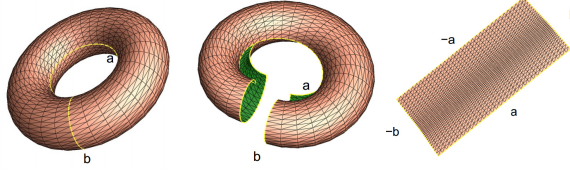


Fig. 6. Left: a and b are 2 homotopy basis for a genus 1 surface. Middle: The surface becomes a topological disk after cut along the its homotopy basis. Right: Flatten the surface onto the plane.

Algorithm 3. Discrete Ricci Flow

Input: Surface M .

Output: The metric U of M .

1. Assign a circle at vertex v_i with radius r_i ; For each edge $[v_i, v_j]$, two circles intersect at an angle ϕ_{ij} , called edge weight.
2. The edge length l_{ij} of $[v_i, v_j]$ is determined by the cosine law: $l_{ij}^2 = r_i^2 + r_j^2 - 2r_i r_j \cos \phi_{ij}$
3. The angle θ_i^{jk} , related to each corner, is determined by the current edge lengths with the inverse hyperbolic cosine law.
4. Compute the discrete Gaussian curvature K_i of each vertex v_i :

$$K_i = \begin{cases} 2\pi - \sum_{f_{ijk} \in F} \theta_i^{jk}, & \text{interior vertex} \\ \pi - \sum_{f_{ijk} \in F} \theta_i^{jk}, & \text{boundary vertex} \end{cases} \quad (2)$$

where θ_i^{jk} represents the corner angle attached to vertex v_i in the face f_{ijk}

5. Update the radius r_i of each vertex v_i : $r_i = r_i - \epsilon K_i r_i$
 6. Repeat the step 2 through 5, until $\|K_i\|$ of all vertices are less than a specific error tolerance.
-

our method and the method in [9], as well as the total running time for de-noise and conformal flattening. Notice that only 1 out of 14 case (3053S) reached genus 0 using method in [9], which means most of the data are not qualified as input of the conformal flattening algorithm, while all 14 reached genus 0 using our de-noise algorithm.

Table 1. De-noise Result and Time Efficiency

Data Index	#Of mesh triangles	# NO. of handles removed by our method	# NO. of handles removed by method in [9]	# Running time (Min.)
3033P	830 K	8	5	8.6 min
3033S	1120 K	26	10	10.3 min
3034P	764 K	24	21	8.1 min
3034S	800 K	7	2	8.2 min
3035S	1060 K	13	6	10.2 min
3036P	875 K	13	8	8.5 min
3037S	836 K	6	0	8.5 min
3038P	1167 K	16	10	10.8 min
3039P	920 K	7	5	9.4 min
3039S	1040 K	9	8	10.0 min
3041P	902 K	15	8	9.7 min
3041S	886 K	4	1	9.1 min
3043P	917 K	11	5	9.4 min
3053S	958 K	4	4	9.6 min

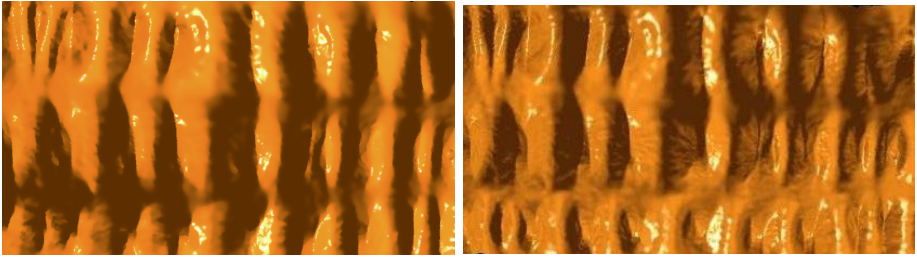


Fig. 7. Left : The zoomed-in results from method in [9]. Right : The zoomed-in results from our method.

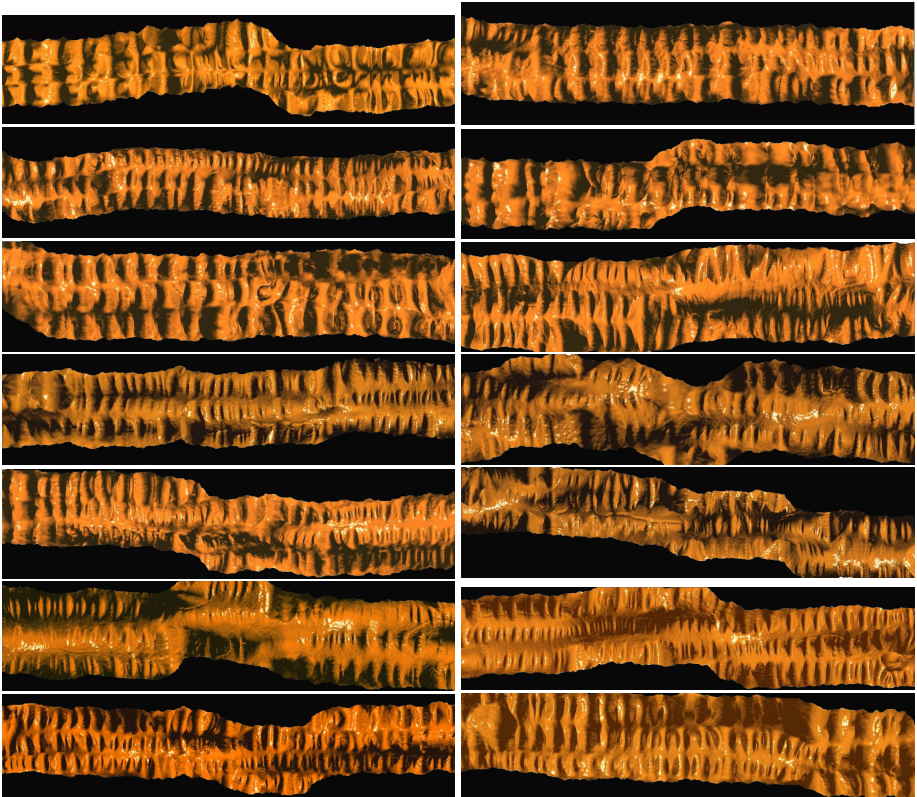


Fig. 8. Zoom-in flatten results of all 14 datasets

For the final colon image, figure 7 shows the zoomed in colon image of [9] and the colon image computed by our method. We can see that the image from our method is clearer and preserved more details compared with the old method. We also show zoom-in flatten results of all 14 datasets in figure 8.

6 Discussion

The key parts of our method is the new efficient topological de-noise algorithm. Our new topological de-noise algorithm guarantee the output to be exactly genus 0. As a result, the whole mapping process becomes much faster and more stable.

References

1. Arbarello, E., Cornalba, M., Griffiths, P., Harris, J.: Topics in the Theory of Algebraic Curves (1938)
2. Balogh, E., Sorantin, E., Nyul, L.G., Palagyi, K., Kuba, A., Werkgartner, G., Spuller, E.: Colon unraveling based on electronic field: Recent progress and future work. In: Proceedings SPIE, pp. 713–721 (2002)
3. Bartrol, A.V., Wegenkittl, R., König, A., Gröller, E., Sorantin, E., Medgraph, T.: Virtual colon flattening (2001)
4. Erickson, J., Whittlesey, K.: Greedy optimal homotopy and homology generators. In: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2005, pp. 1038–1046 (2005)
5. Haker, S., Angenent, S., Kikinis, R.: Nondistorting flattening maps and the 3d visualization of colon ct images. *IEEE Trans. on Medical Imaging* 19, 665–670 (2000)
6. Hamilton, R.S.: The Ricci flow on surfaces. Mathematics and general relativity (Santa Cruz, CA, 1986). *Contemp. Math. Amer. Math. Soc. Providence, RI* 71 (1988)
7. Hong, L., Kaufman, A., Wei, Y., Viswambharan, A., Wax, M., Liang, Z.: 3d virtual colonoscopy. In: IEEE Symposium on Frontier in Biomedical Visualization, pp. 26–32 (1995)
8. Hong, L., Liang, Z., Viswambharan, A., Kaufman, A., Wax, M.: Reconstruction and visualization of 3d models of the colonic surface. *IEEE Transactions on Nuclear Science*, 1297–1302 (1997)
9. Hong, W., Gu, X., Qiu, F., Jin, M., Kaufman, A.: Conformal virtual colon flattening. In: ACM Symposium on Solid and Physical Modeling, pp. 85–94 (2006)
10. Hong, W., Qiu, F., Kaufman, A.: A pipeline for computer aided polyp detection. *IEEE Transactions on Visualization and Computer Graphics* 12, 861–868 (2006)
11. Jin, M., Kim, J., Gu, X.D.: Discrete surface ricci flow: Theory and applications. In: IMA Conference on the Mathematics of Surfaces, pp. 209–232 (2007)
12. Massey, W.: *Algebraic Topology: An Introduction*. Springer (1967)
13. Paik, D., Beaulieu, C., Jeffrey, R., Karadi, C.A., Napel, S.: Visualization modes for ct colonography using cylindrical and planar map projections. *Journal of Computer Assisted Tomography*, 179–188 (2000)
14. Pickhardt, P.J., Choi, J.R., Hwang, I., Butler, J.A., Puckett, M.L., Hildebrandt, H.A., Wong, R.K., Nugent, P.A., Mysliwiec, P.A., Schindler, W.R.: Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *New England Journal of Medicine* 349(23), 2191–2200 (2003)
15. Hong, W., Qiu, F., Kaufman, A.: A pipeline for computer aided polyp detection. *IEEE Transactions on Visualization and Computer Graphics*, 861–868 (2006)
16. Luboldt, W., Steiner, P., Bauerfeind, P., Pelkonen, P., Debatin, J.: Detection of mass lesions with mr colonoscopy: Preliminary report. *Radiology* 207, 59–65 (1998)

17. Wang, S., Li, L., Cohen, H., Mankes, S., Chen, J., Liang, Z.: An em approach to map solution of segmenting tissue mixture percentages with application to ct-based virtual colonoscopy. *IEEE Transactions on Medical Imaging* 28, 297–310 (2009)
18. Wang, Z., Li, B., Liang, Z.: Feature-based texture display for detection of colonic polyps on flattened colon volume. In: *Intl. Conf. of IEEE Engineering in Medicine and Biology* (2005)
19. Zeng, W., Samaras, D., Gu, X.D.: Ricci flow for 3D shape analysis. *PAMI* 32(4), 662–677 (2010)

Adaptive Volumetric Detection of Lesions for Minimal-Preparation Dual-Energy CT Colonography

Janne J. Näppi¹, Se Hyung Kim²,
and Hiroyuki Yoshida¹

¹ 3D Imaging Research, Department of Radiology,
Massachusetts General Hospital and Harvard Medical School,
25 New Chardon St., Suite 400C, Boston, Massachusetts 02114, USA
jnappi@partners.org, Hiroyuki.Yoshida@mgh.harvard.edu
² Seoul National University Hospital, 101 Daehangno,
Chongno-gu, Seoul, 110-744, Republic of Korea

Abstract. Dual-energy computed tomographic colonography (DE-CTC) provides detailed information about the chemical composition of colon that can be used to improve the accuracy of computer-aided detection (CAD). We investigated how to calculate a thick target region for volumetric detection of lesions in DE-CTC. After automated extraction of the region of colonic lumen, the target region is calculated by use of a distance-based scheme, where the image scale of the shape features that are used for the detection of lesion candidates is adapted to the thickness of the target region. False-positive (FP) detections are reduced by use of a random-forest classifier. The detection accuracy of the CAD scheme was evaluated at 5 thicknesses of the target region by use of a leave-one-patient-out evaluation with 23 clinical minimal-preparation DE-CTC cases including 27 lesions ≥ 6 mm in size. The results indicate that the optimal choice of thickness depends on the size and morphology of the target lesion. At optimal thickness, the per-patient sensitivity was 100% at 5 FP detections per patient on average, where the per-lesion sensitivity was 100% (94%) for lesions ≥ 10 mm (6 – 9 mm) in size. The results compare favorably with those of our previous approach.

Keywords: Computer-aided detection, dual energy, polyp detection, laxative-free, non-cathartic, virtual colonoscopy, computed tomographic colonography.

1 Introduction

Colorectal cancer is one of the leading causes of cancer mortality in Western countries, but it could be prevented by early detection and removal of the precursor lesions [24]. Computed tomographic colonography (CTC) is a minimally invasive imaging procedure that can detect patients with large adenomas and cancers at high sensitivity [5].

Computer-aided detection (CAD), that indicates locations of lesions in CTC images automatically to radiologists, can be used to increase the detection sensitivity and consistency of CTC examinations [2,26]. Most CAD schemes have been developed for conventional CTC, where the patients are prepared with full cathartic bowel cleansing. Most of these CAD schemes tend to miss many small polyps (6 – 9 mm in size) [8] and flat lesions [12], and they display many false-positive (FP) detections [20]. Recent developments have focused on improving the detection sensitivity for polyps 6 – 9 mm in size. In one study, an advanced interpolation method improved the detection sensitivity of CAD from 55% to 66% at 10 FP detections per patient [10]. In another study, multi-objective optimization improved the detection sensitivity from 62% to 75% at 10 FP detections per patient.

Most CAD schemes appear to use a surface-based detection scheme that analyzes curvatures of a triangulated surface mesh of the colon. Previously, we developed a CAD scheme where lesions are detected within a thick volumetric target region encompassing the colonic mucosa [25,15]. In a recent evaluation with 1,948 patients from 32 medical institutions [22], the CAD scheme detected 82% of the polyps 6 – 9 mm at 12 FP detections per patient and 100% (81%) of the flat lesions ≥ 10 mm (6 – 9 mm) in size at 5 FP detections per patient. Importantly, the difference in the detection accuracy between polyps and flat lesions was not statistically significant. Although the number of FP detections is still relatively high, these results indicate that the volumetric detection has significant advantages over the conventional surface-based detection scheme.

One of the most common sources of FP CAD detections are partial-volume tagging artifacts that imitate the shape and density of lesions [20]. Such artifacts can be reduced by the application of dual-energy CTC (DE-CTC) [14] that uses two monochromatic x-ray beams to provide more information about the chemical composition of colon than what is available with conventional single-energy CTC [6]. This can be useful in implementing non-cathartic CTC examinations for increasing patient adherence to colorectal screening guidelines [1].

Previously, we observed that the thickness of the volumetric target region can have significant effect on the detection accuracy of CAD in DE-CTC [19]. Therefore, in this study, we developed a novel adaptive volumetric detection scheme for DE-CTC. The target region is extracted precisely at specified thickness by use of a distance-based method, and the calculation of shape features is adapted to the thickness of the target region. To determine an optimal wall-thickness value, the detection accuracy of CAD was evaluated at several thicknesses of the target region by use of clinical DE-CTC cases.

2 Methods

2.1 Materials

Twenty-five patients were prepared for a DE-CTC examination by use of a one-day preparation. In the evening prior to the examination, the patients were advised to consume LoSo Prep (E-Z-EM, Inc., New York, USA) with several

cups of water. In the morning of the examination, the patients consumed 50 ml of iodine for positive-contrast tagging of residual bowel contents. The DE-CTC scans were acquired in supine and prone positions with a dual-energy CT scanner (SOMATOM Definition, Siemens Healthcare) at 140 kVp and 80 kVp energies by use of a 1-mm slice thickness. No intravenous contrast was used. After the CTC acquisition, the patients underwent conventional colonoscopy. Expert radiologists correlated the DE-CTC data with the findings of colonoscopy.

All 25 DE-CTC cases were included in the study regardless of their diagnostic quality. They were divided randomly into a development set of 2 cases and an independent evaluation set of 23 cases. The development set was used for parameter estimation, whereas the evaluation set was used for assessing the detection accuracy of the CAD scheme.

2.2 Extraction of the Colonic Lumen

To extract the region of colon from DE-CTC image data, a linearly mixed volume is calculated from the 140 kVp and 80 kVp energy images for each input CTC series. The mixed volume is used for providing shape information, whereas the two energy images are used for the characterization of underlying materials. The region of colonic lumen is extracted by use of a fully automated lumen-based tracking method [17]. Dual-energy index (DEI) is calculated to delineate fecal-tagged materials precisely from soft tissue as

$$DEI = \frac{v_{80} - v_{140}}{2000 + v_{80} + v_{140}}, \quad (1)$$

where v_{140} and v_{80} are the dual-energy values. The application of the DEI is explained in Fig. 1.

2.3 Distance-Based Extraction of the Target Region

Let L denote the extracted three-dimensional binary mask of the colonic lumen (Fig. 2a). To extract a thick region encompassing the colonic wall, the surface of L is adjusted to an iso-intensity CT value level of -500 Hounsfield units (HU) — this value has been established as an optimal value for representing the surface of colon in CTC [23]. In tagged regions, the surface of L is adjusted to a pseudo-enhancement-corrected [18] CT value of 150 HU.

Let L' denote the adjusted binary lumen mask. Suppose that the target region encompassing the colonic surface should have a thickness of d millimeters (mm). To do this, we calculate a signed three-dimensional distance transform of L' , or $D_{L'}$ [3] (Fig. 2b). Negative values of $D_{L'}$ indicate distance to the lumen surface inside the lumen, whereas positive values indicate distance to the lumen surface outside of the lumen. Therefore, we can establish the final target region by clipping of the distance values between $-d/2$ and $d/2$ mm (Fig. 2c).

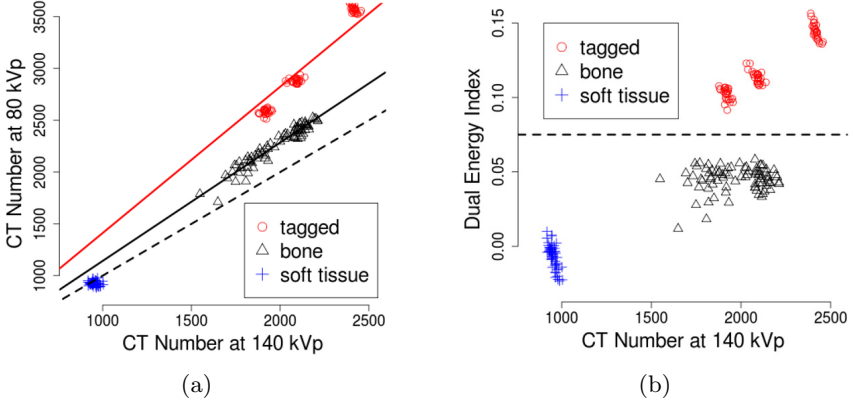


Fig. 1. (a) When the CT numbers of the 140 kVp and 80 kVp energy images are mapped to a plane, samples of materials with different effective atomic numbers (Z) appear along their unique characteristic lines. Water-like materials ($Z = 7$) and soft-tissue materials ($Z < 10$) appear along the line of identity (dotted line), whereas heavier bone-like materials ($Z = 14$) and fecal-tagged residual materials (iodine; $Z=56$) appear above the line of identity. (b) Fecal-tagged materials can be delineated by thresholding (dotted line) of the dual-energy index values.

2.4 Adaptive Volumetric Detection of Lesions

To detect lesions, we calculate two shape features at each point of the extracted target region: the volumetric shape index (SI) and curvedness (CV). They can be expressed as $SI(p) = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k_1(p)+k_2(p)}{k_1(p)-k_2(p)}$ and $CV(p) = \sqrt{\frac{k_1(p)^2+k_2(p)^2}{2}}$, where $k_1(p)$ and $k_2(p)$ are the principal curvatures of a surface passing through a voxel p [7]. Because $k_1(p)$ and $k_2(p)$ can be calculated implicitly from the three-dimensional first- and second-order partial derivatives of CT values [11], the SI and CV are defined at every voxel of a CTC volume.

The SI maps every distinct shape (except for plane) to a unique value. The highest values of SI indicate shapes of colonic lesions, whereas lower values indicate other colonic structures [25] (Fig. 21). The SI is not affected by the flatness or sharpness of the geometric shape. Therefore, a high value of SI can indicate both polyps and flat lesions. The perceived flatness or sharpness of the shape is indicated by the CV : a low value of CV indicates a flat-topped morphology, whereas a high value indicates a sharp-crested morphology.

Before the shape calculations, the CTC volume is convolved with a three-dimensional Gaussian kernel,

$$G_3(x, y, z, \sigma) = \frac{1}{\sqrt{2\pi}\sigma^3} \exp^{-(x^2+y^2+z^2)/(2\sigma^2)}, \quad (2)$$

to obtain an infinitely differentiable function [13]. This will also delimit the image scale of observable lesions to the width of the Gaussian, or σ [9]. Because

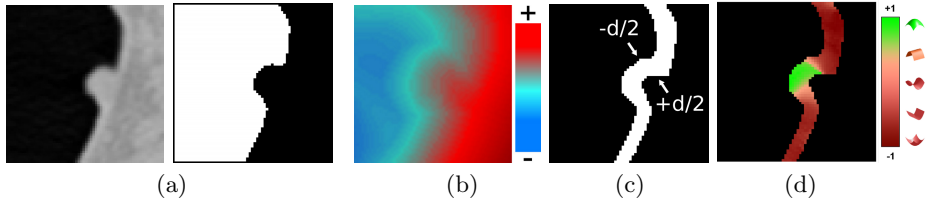


Fig. 2. Extraction of a thick target region encompassing the colonic wall (the actual calculations are performed three-dimensionally). (a) Left: axial CTC image of a region of interest with a polyp. Right: the lumen-based tracking method extracts a binary lumen mask (white region). (b) Signed distance transform of the lumen mask: colors indicate distance to lumen surface. (c) A target region of thickness d (white region) can be extracted by clipping of distance values between $-d/2$ and $d/2$. (d) To detect lesions, shape features (here: shape index) are calculated for voxels of the target region.

clinically significant lesions are ≥ 6 mm in largest diameter, it might seem appropriate to use a convolution kernel with $\sigma = 6$ mm. However, the optimal choice of σ is not clear, because convolution with a wide Gaussian can distort important details, thereby making subtle lesions challenging to detect.

In a previous study, where the shape features were calculated at a fixed image scale of $\sigma = 4$ mm, the optimal thickness of the target region was 4 mm. This suggests that the optimal image scale of features might depend on the thickness of the target region. Therefore, in this study, we considered adaptive detection where the image scale (σ) of the SI and CV features is set to the thickness (d) of the target region (Fig. 3).

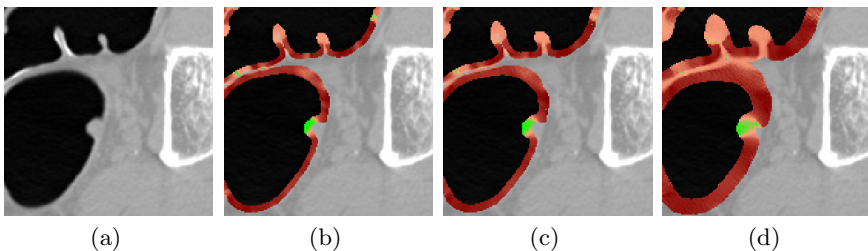


Fig. 3. (a) Axial CTC image of a region of interest with a polyp at center. (b) The values of shape index overlaid on a relatively thin target region (see Fig. 2 for explanation of the color coding). (c) Adjustment of scale to a relatively thick target region. (d) Adjustment of scale to a thick target region.

After the calculation of shape features, lesion candidates are detected by hysteresis thresholding of the shape values [25]. Smallest detections are excluded to avoid the detection of clinically insignificant lesions (those measuring < 6 mm in largest diameter).

2.5 Reduction of False Positives

The detection step yields a large number of lesion candidates, many of which are FP detections. To minimize the number of false positives, a number of shape, texture, and dual-energy features are calculated for the regions of lesion candidates [16]. The dual-energy features include the DEI (Equation (1)), dual-energy ratio $DER = \frac{v_{80}}{1+v_{140}}$, and dual-energy mixtures $DEM_r = rv_{140} + (1-r)v_{80}$ ($r \in \{0.33, 0.50, 0.66\}$). Fig. 4 shows an example of the application of the DEI.

For analysis, the per-voxel values of features are summarized into aggregate features in terms of their mean and variance within the region of a lesion candidate. The lesion candidates are classified into true-positive and FP detections by use of a random-forest classifier [4] that has been previously shown to outperform support vector machines significantly in the detection of colorectal lesions in CTC [21]. The classifier calculates the likelihood that a lesion candidate represents a true lesion. The lesion candidates with highest likelihood of being true lesions are displayed as the lesions detected by CAD.

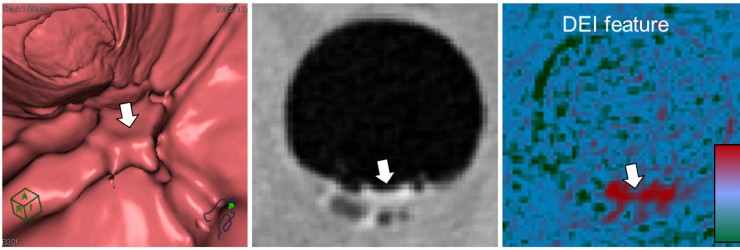


Fig. 4. Left: virtual endoluminal view of an obscure object (indicated by arrow) on fold. Middle: axial CTC image of the object. Right: high values of the dual-energy index in the axial image indicate that the object is a partial-volume tagging artifact.

2.6 Evaluation

The detection accuracy of CAD was evaluated at wall-thickness values of 2 mm, 3 mm, 4 mm, 5 mm, and 6 mm, by use of a leave-one-patient-out evaluation. A colonoscopy-confirmed lesion was considered to be detected correctly by CAD, if the center of a lesion candidate detected by CAD was within the radius of the center of a true lesion in CTC data. All other CAD detections were considered as false positives. To simulate a clinically realistic application of CAD, the number of displayed detections was limited to at most 15 per patient.

3 Results

There were 11 lesions ≥ 10 mm and 16 lesions 6 – 9 mm in largest diameter in the 23 patients. These included 4 cancers, 20 adenomas, and 3 hyperplastic lesions.

Fig. 5 shows the effect of the thickness of target region on the per-lesion sensitivity of the CAD scheme. For lesions ≥ 10 mm in size (solid line), the detection sensitivity was 100% at all thicknesses except at $d = 2$ mm. For lesions 6 – 9 mm in size (dotted line), the detection sensitivity peaked out at $d = 3$ mm (94%), and then it declined rapidly as the wall thickness increased. At $d = 6$ mm, the reduction in detection sensitivity was statistically significant ($p < 0.05$). Regardless of the wall thickness, all four carcinomas were detected correctly.

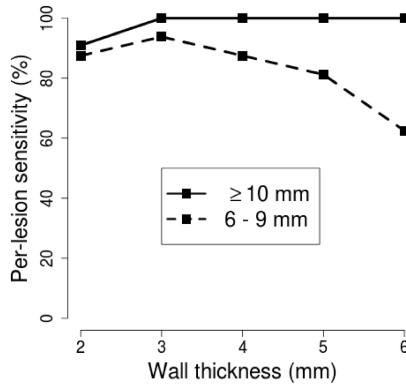


Fig. 5. Per-lesion sensitivity of CAD in terms of the thickness of target region

Changes in wall thickness affect not only the sensitivity but also specificity of the detection. Fig. 6 illustrates the CAD performance at different thicknesses of the target region. The numbers in the plot indicate wall-thickness values (in mm), and their location in the plot indicates the per-lesion sensitivity and the average number of FP detections per patient. The highest detection performance would be located at the top left corner of the plot, whereas the lowest detection performance would be located at the bottom right corner. For lesions ≥ 10 mm in size, the detection performance was highest at a wall thickness of 4 mm — at higher wall-thickness values, the detection specificity was reduced due to a 13-mm flat lesion that was not detected anymore because of the effects of image smoothing. For lesions 6 – 9 mm in size, at $d = 3$ mm, 94% of the polyps were detected at 11 FPs per patient. Only one lesion, a 7-mm sessile polyp, was missed, and our retrospective analysis indicates that this happened because the polyp appears to measure only 4 mm in the CTC images.

Finally, Fig. 7 presents the free-response receiver operating characteristic curves of the per-lesion and per-patient detection accuracy of the CAD scheme at the 3-mm wall-thickness value. At 5 FP detections per patient on average, all patients with lesions ≥ 6 mm were detected correctly.

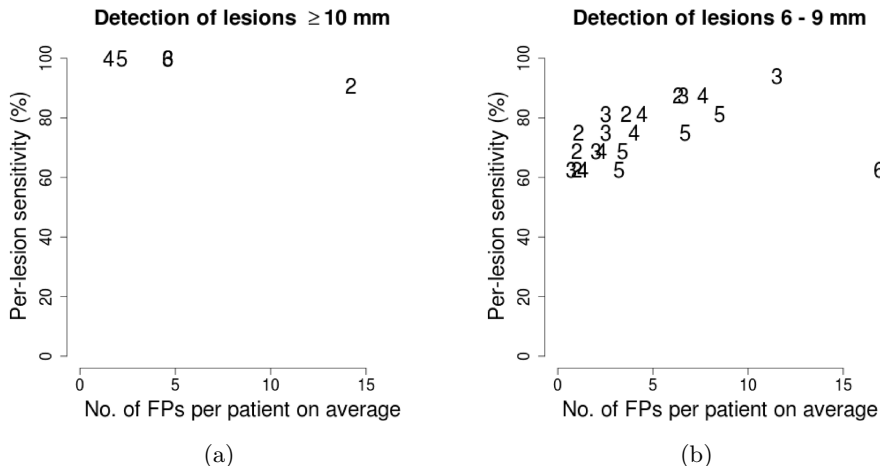


Fig. 6. The numbers on the plot indicate CAD performance at that thickness of the target region (in mm). (a) Detection of lesions ≥ 10 mm in size. (b) Detection of lesions 6 – 9 mm in size.

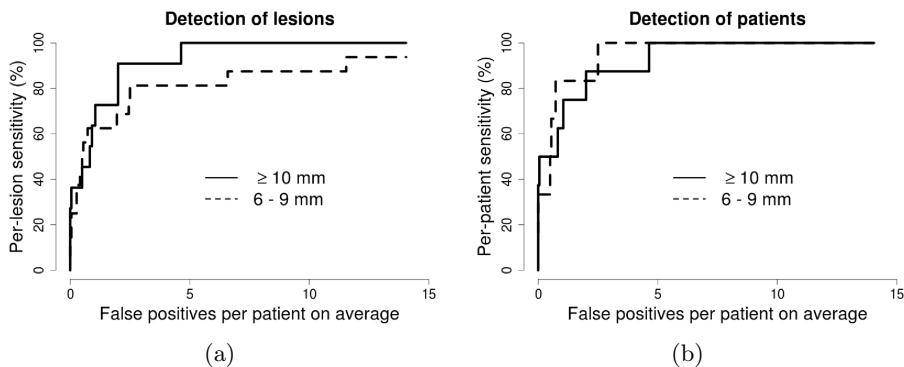


Fig. 7. Free-response receiver operating characteristic curves of the (a) per-lesion and (b) per-patient detection accuracy of the CAD scheme for lesions ≥ 10 mm and 6 – 9 mm in size at 3-mm thickness of the target region

4 Discussion

We developed a novel method for volumetric detection of lesions in DE-CTC. A thick volumetric target region encompassing the colonic wall is extracted from CTC images by use of a distance-based scheme, and the scale of the shape detection features is adapted to the thickness of the target region.

The results indicate that the thickness of the target region has a significant effect on the detection accuracy of CAD. The effect is most noticeable in the detection of 6 – 9 mm lesions. For larger lesions, the wall thickness has less effect on detection sensitivity, but it can affect the detection specificity.

In a previous study, where the shape features were calculated at a fixed image scale using the same clinical DE-CTC cases [19], the per-lesion detection sensitivity for 6 – 9 mm lesions was 87.5% with 11 FP detections per patient at a wall thickness of $d = 4$ mm. In this study, the use of adaptive volumetric detection at a wall thickness of 3 mm yielded a higher detection sensitivity of 94%. Furthermore, at a comparable 87.5% sensitivity between the studies, in this study the number of FP detections was reduced by 45%, or to 6 per patient. Therefore, the use of the adaptive volumetric detection method of this study compares favorably with our previous approach.

Acknowledgements. This work was supported in part by grants of Prevent Cancer Foundation, NIH R21CA140934, CA095279, CA166816, and CA131718. We also thank Partners Research Computing for providing high-performance computing services.

References

1. Beebe, T., Johnson, C., Stoner, S., Anderson, K., Limburg, P.: Assessing attitudes toward laxative preparation in colorectal cancer screening and effects on future testing: potential receptivity to computed tomographic colonography. *Mayo. Clinic. Proc.* 82, 666–671 (2007)
2. Bielen, D., Kiss, G.: Computer-aided detection for CT colonography: update 2007. *Abdom. Imaging* 32, 571–581 (2007)
3. Borgefors, G.: On digital distance transforms in three dimensions. *Comput. Vis. Image Und.* 64, 368–376 (1996)
4. Breiman, L.: Random forests. *Mach. Learn.* 45, 5–32 (2001)
5. Johnson, C., Chen, M.H., Toleday, A., Heiken, J., Dachman, A., Kuo, M., et al.: Accuracy of CT colonography for detection of large adenomas and cancers. *N. Engl. J. Med.* 359, 1207–1217 (2008)
6. Johnson, T., Krauss, B., Seldmair, M., Grasruck, M., et al.: Material differentiation by dual energy CT: initial experience. *Eur. Radiol.* 17, 1510–1517 (2006)
7. Koenderink, J., Doorn, R.: Surface shape and curvature scales. *Image and Vision Computing* 10, 557–565 (1992)
8. Lawrence, E., Pickhardt, P., Kim, D., Robbins, J.: Colorectal polyps: stand-alone performance of computer-aided detection in a large asymptomatic screening population. *Radiology* 256, 791–798 (2012)
9. Lindeberg, T.: *Scale-scape theory in computer vision*. Kluwer Academic Publishers, Boston (1993)
10. Liu, J., Kabadi, S., Van Uitert, R., Petrick, N., Deriche, R., Summers, R.: Improved computer-aided detection of small polyps in CT colonography using interpolation for curvature estimation. *Med. Phys.* 38, 4276–4284 (2011)
11. Lohmann, G.: *Volumetric image analysis*. John Wiley & Sons, New York (1998)
12. Mang, T., Bogoni, L., Salganicoff, M., Wolf, M., Raykar, V., Macari, M., et al.: Computer-aided detection of colorectal polyps in CT colonography with and without fecal tagging: a stand-alone evaluation. *Invest. Radiol.* 47, 99–108 (2012)
13. Näppi, J., Frimmel, H., Yoshida, H.: Virtual endoscopic visualization of the colon by shape-scale signatures. *IEEE Trans. Inf. Technol. Biomed.* 9, 120–131 (2005)

14. Näppi, J., Kim, S., Yoshida, H.: Automated detection of colorectal lesions with dual-energy CT colonography. In: van Ginneken, B., Noval, C. (eds.) *SPIE Medical Imaging 2012: Computed-Aided Diagnosis*, vol. 8315, pp. 83150Y1–83150Y6 (2012)
15. Näppi, J., MacEneaney, P., Dachman, A., Yoshida, H.: Knowledge-guided automated segmentation of colon for computer-aided detection of polyps in CT colonography. *J. Comput. Assist. Tomogr.* 26, 493–504 (2002)
16. Näppi, J., Yoshida, H.: Automated detection of polyps in CT colonography: evaluation of volumetric features for reduction of false positives. *Acad. Radiol.* 9, 386–397 (2002)
17. Näppi, J., Yoshida, H.: Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography. *Acad. Radiol.* 25, 287–300 (2007)
18. Näppi, J., Yoshida, H.: Adaptive correction of the pseudo-enhancement of CT attenuation for fecal-tagging CT colonography. *Med. Image Anal.* 12, 413–426 (2008)
19. Näppi, J., Kim, S., Yoshida, H.: Volumetric detection of colorectal lesions for non-cathartic dual-energy computed tomographic colonography. In: *34th Annual International Conference of the IEEE Engineering in Medicine & Biology Society*. IEEE Press, San Diego (in press, 2012)
20. Näppi, J., Nagata, K.: Sources of false positives in computer-assisted CT colonography. *Abdom. Imaging* 36, 153–164 (2011)
21. Näppi, J.J., Regge, D., Yoshida, H.: Comparative Performance of Random Forest and Support Vector Machine Classifiers for Detection of Colorectal Lesions in CT Colonography. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) *Abdominal Imaging 2011*. LNCS, vol. 7029, pp. 27–34. Springer, Heidelberg (2012)
22. Näppi, J., Yoshida, H., Regge, D.: Accuracy of computer-aided detection in detecting flat lesions in CT colonography. *Insights into Imaging* 3(S1), S334 (2012)
23. Park, S., Choi, E., Lee, S., Woo, J., et al.: Linear polyp measurement at CT colonography: 3D endoluminal measurement with optimized surface-rendering threshold value and automated measurement. *Radiology* 246, 157–167 (2008)
24. Winawer, S., Fletcher, R., Miller, L., Godlee, F., Stolar, M., Mulrow, C., Woolf, S., Glick, S., Ganiats, T., Bond, J., Rosen, L., Zapka, J., Olsen, S., Giardiello, F., Sisk, J., Van Antwerp, R., Brown-Davis, C., Marciniak, D., Mayer, R.: Colorectal cancer screening: clinical guidelines and rationale. *Gastroenterology* 112, 594–642 (1997)
25. Yoshida, H., Näppi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *IEEE Trans. Med. Imaging* 20, 1261–1274 (2001)
26. Yoshida, H., Näppi, J.: CAD in CT colonography without and with oral contrast agents: progress and challenges. *Comp. Med. Imag. Graph* 31, 267–284 (2007)

Computer-Aided Detection for Ultra-Low-Dose CT Colonography

Janne J. Näppi¹, Masanori Imuta²,
Yasuyuki Yamashita², and Hiroyuki Yoshida¹

¹ 3D Imaging Research, Department of Radiology,
Massachusetts General Hospital and Harvard Medical School,
25 New Chardon St., Suite 400C, Boston, Massachusetts 02114, USA
jnappi@partners.org, yoshida.hiro@mg.harvard.edu

² Kumamoto University, Department of Diagnostic Radiology,
Kumamoto 860-8555, Japan

Abstract. To screen large populations for colorectal cancer, it may be necessary to reduce the radiation dose of computed tomographic colonography (CTC) examinations. We compared the accuracy of computer-aided detection (CAD) in standard-dose (SD) CTC with that in ultra-low-dose (ULD) CTC. We also assessed the effect of linear and nonlinear denoising methods on CAD performance in ULD CTC. The CAD system was trained to detect polyps with 43 SD CTC studies. It was tested with 24 clinical studies, where the supine series were acquired with SD CTC and the prone series were acquired with ULD CTC. The polyp detection accuracy of CAD was significantly lower in ULD CTC than in SD CTC. Linear denoising of ULD CTC images improved the detection accuracy for large polyps, but it reduced sensitivity for small polyps. However, with nonlinear denoising, the detection accuracy of CAD in ULD CTC was not significantly different from that in SD CTC.

Keywords: Computer-aided detection, dose, diffusion, polyp detection, virtual colonoscopy, computed tomographic colonography.

1 Introduction

Colorectal cancer is one of the leading causes of cancer mortality. Studies have shown that the per-patient sensitivity of computed tomographic colonography (CTC) for large adenomas and cancers is comparable to that of conventional colonoscopy [5,16]. This makes CTC a potentially useful technique for colorectal screening [8].

However, if CTC was used in large-scale population screening, its target population would be >100 million people in the United States alone [1]. One concern is the potential risk of radiation-induced cancers. The risk is theoretical, because at low doses there are many uncertainties with regard to the true effect of ionizing radiation [11]. Nevertheless, some simulation studies have suggested that, if performed every 5 years for ages 50 – 80, CTC could induce up to 60,000 cancers per 100 million screened individuals [4].

Tube current is one of the principal sources of ionizing radiation in CTC. Studies have suggested that a reduction in tube current might not have a significant effect on polyp detection in CTC [3]. Therefore, it may be possible to perform ultra-low-dose (ULD) CTC. The radiation dose of 1 – 2 millisievert (mSv) of an ULD CTC examination would compare favorably with the yearly dose of 2.5 – 3.0 mSv of the normal background radiation exposure in the United States [18]. However, the use of ULD CTC can distort the observed colon surface and lesions (Fig. 1), and this can be a problem for the application of CAD.

To date, the effect of ULD CTC on CAD has been investigated by use of simulations only. A phantom study suggested that polyps ≥ 6 mm in size can be detected at an effective dose of 1.61 mSv (a current of 5 – 15 mAs), if the CTC data are preprocessed by use of an anisotropic filter [10]. A study based on pig colons indicated that tube currents of 10 mAs are not an influencing factor of CAD performance [9]. In another simulation study, where image noise was added to 23 clinical studies to imitate an effective dose of 0.17 mSv (a current of 1.39 mAs), conventional CAD detected 78% of polyps ≥ 10 mm and 65% of polyps 6 – 9 mm in size at 16 false-positive (FP) detections per patient [17].

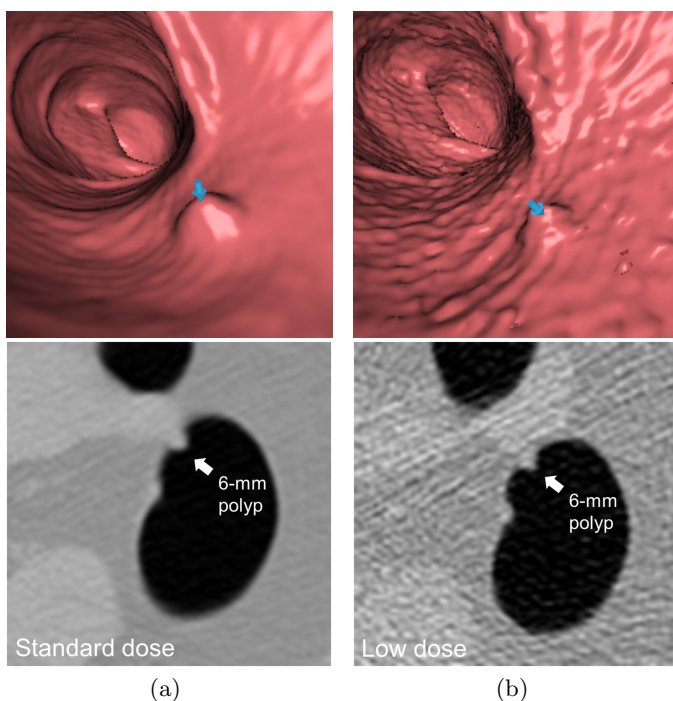


Fig. 1. Visualization of a 6-mm polyp (indicated by arrow) in (a) standard-dose and (b) ultra-low-dose CT colonography

In this study, we investigated the effect of ULD CTC on the detection accuracy of CAD by use of clinical standard dose (SD) and ULD CTC studies. We also assessed the effect of linear and nonlinear denoising algorithms on CAD performance in ULD CTC.

2 Methods

2.1 Gaussian Filtering

Gaussian filtering is a linear method that is often used for reducing image noise (Fig. 2b). Let I denote a volumetric input CTC image. The convolution of I with a three-dimensional (3D) Gaussian kernel can be written as

$$I \cdot G_3(x, y, z, \sigma) = \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} I(x, y, z) G_3(x - i, y - j, z - k, \sigma), \quad (1)$$

where the 3D Gaussian kernel is of the form

$$G_3(x, y, z, \sigma) = \frac{1}{\sqrt{2\pi}\sigma^3} \exp^{-(x^2+y^2+z^2)/(2\sigma^2)}. \quad (2)$$

The parameter σ determines the width of the Gaussian kernel.

2.2 Diffusion Filtering

Diffusion filtering extends the concept of the Gaussian filtering by reducing the effect of smoothing at image boundaries (Fig. 2b). Such a nonlinear filter preserves surfaces of the image more faithfully than conventional Gaussian filtering.

The anisotropic diffusion equation can be written as [15]

$$\frac{\partial I}{\partial t} = \text{div}(c(x, y, z, t)\nabla I) = \nabla c \cdot \nabla I + c(x, y, z, t)\Delta I, \quad (3)$$

where $c(x, y, t)$ is a diffusion coefficient that controls the rate of the diffusion. Suppose that we know the locations of image boundaries, and that we would like to encourage smoothing within homogeneous regions in preference to smoothing of boundaries. To approximate such a function, we can define

$$c(\|\nabla I\|) = \frac{1}{1 + (\frac{\|\nabla I\|}{K})^2}, \quad (4)$$

where K controls sensitivity to image boundaries.

Fig. 3 illustrates the outcomes of Gaussian filtering and diffusion filtering. When the original image (Fig. 3a) is subjected to Gaussian filtering (Fig. 3b), the perceived image noise is reduced, but also the boundaries and the polyp at the center of the image are blurred. When the original image is subjected to diffusion filtering (Fig. 3c), the boundaries and the polyp are preserved better than after Gaussian filtering.

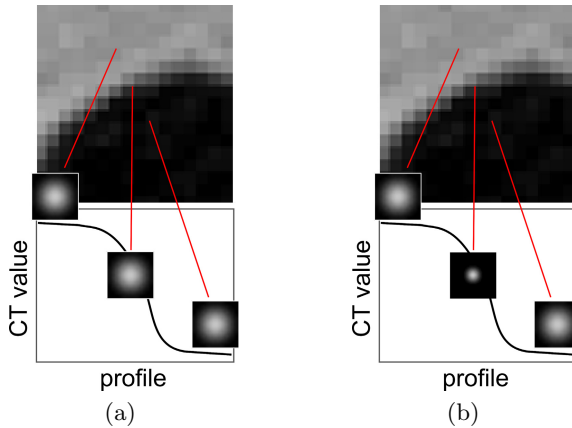


Fig. 2. Illustration of the denoising algorithms. The lines indicate matching locations between the CTC image of a colonic wall (top image) and an intensity profile of the region (bottom image). (a) In Gaussian filtering, the same smoothing kernel is applied everywhere in the image. (b) In diffusion filtering, the width of the smoothing kernel varies according to the intensity changes of the underlying image.

2.3 The CAD Scheme

Fig. 4 shows an overview of the CAD scheme. The region of colonic lumen is extracted automatically [14]. Shape-based polyp detection is performed within a thick region encompassing the colonic wall [19]. After the calculation of shape and texture features of detected regions [13], a Bayesian neural network is used to establish a lesion likelihood for each detection [12]. The detections with highest lesion likelihood are displayed to the user as the output of the CAD scheme.

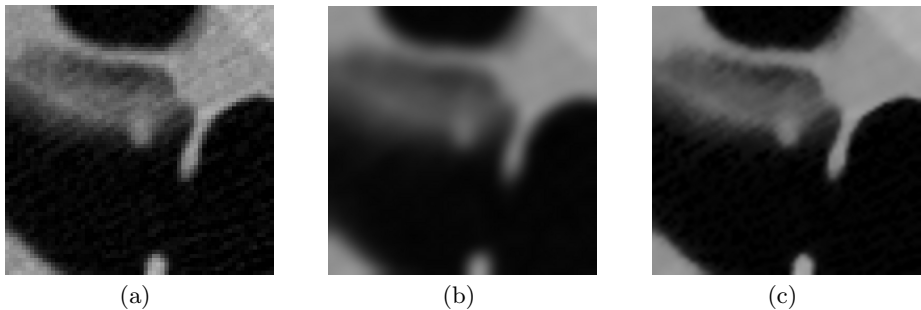


Fig. 3. (a) Axial CTC image with a polyp at center. (b) Convolution with a Gaussian filter. (c) Convolution with a diffusion filter.

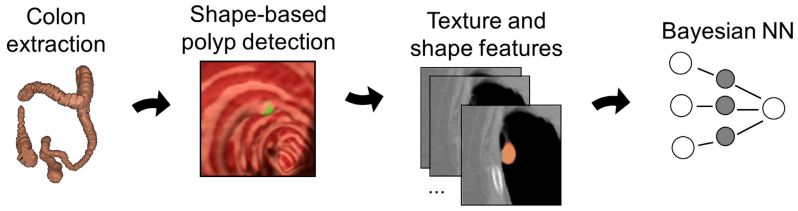


Fig. 4. Overview of the CAD scheme

2.4 Evaluation

Sixty-seven patients were prepared for a CTC examination with a cathartic pre-colonoscopy bowel cleansing. The CTC examinations were performed in supine and prone positions with a 2.5-mm slice thickness and a 1.25-mm reconstruction interval. After a CTC examination, conventional colonoscopy was performed to confirm the detected lesions.

The patients were divided into a training regimen and a testing regimen. The training regimen contained 43 patients (18 males, 25 females; mean age: 61 years). The patients were scanned at two hospitals with SD CTC (100 mA current and 120 kVp voltage). There were 44 lesions measuring ≥ 6 mm in largest diameter in 24 patients (Fig. 5a).

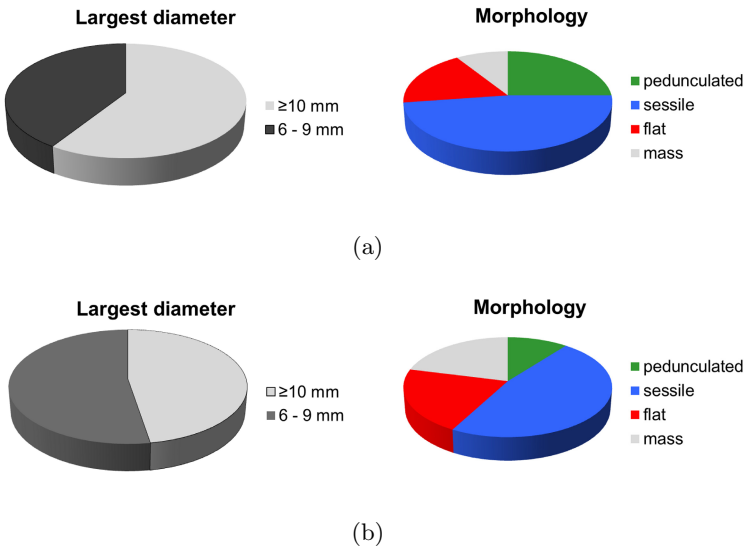


Fig. 5. Distribution of lesions in (a) the training regimen and (b) the testing regimen

The testing regimen contained 24 patients (17 males, 7 females; mean age: 63 years). The supine CTC series were acquired with SD CTC (100 mA current and 120 kVp voltage), whereas the prone CTC series were acquired with ULD CTC (20 mA current and 120 kVp voltage). There were 19 lesions ≥ 6 mm in 12 patients (Fig. 5b).

The CAD scheme was trained to detect lesions with the training regimen. It was tested with the testing regimen by use of four data sets: (1) the 24 supine SD CTC series, (2) the 24 prone ULD CTC series, (3) the 24 prone ULD CTC series with Gaussian filtering, and (4) the 24 prone ULD CTC series with diffusion filtering. Statistical significance of the difference between detection accuracies was tested by use of the jack-knife free-response receiver operating characteristic analysis [2].

3 Results

The training sensitivity of CAD was 93% for lesions ≥ 6 mm in size at 5 FP detections per patient. The sensitivity was 100% for lesions ≥ 10 mm and 83% for lesions 6 – 9 mm in size.

Table 1 shows the per-lesion detection accuracy of CAD on the testing data. With the SD series, the CAD scheme detected 100% of lesions ≥ 10 mm and 90% of lesions 6 – 9 mm in size at 5 FPs per patient. With the ULD series (without denoising), the detection sensitivity decreased significantly for lesions ≥ 10 mm in size ($p < 0.05$), for lesions 6 – 9 mm in size ($p < 0.05$), and for lesions ≥ 6 mm in size ($p < 0.05$).

Table 1. Per-lesion accuracy of CAD on the testing set. Arrows indicate improvement (\uparrow) or reduction (\downarrow) of detection sensitivity by Gaussian filtering (G) or by diffusion filtering (D). Abbreviations: ULD = ultra-low-dose, FPs/study = median number of false-positive detections per patient.

Dose	≥ 10 mm	6 – 9 mm	≥ 6 mm	FPs/study
Standard	100%	90%	95%	5
ULD	78%	70%	74%	4
ULD+G	\uparrow 89%	\downarrow 60%	74%	5
ULD+D	\uparrow 100%	70%	\uparrow 84%	5

When Gaussian denoising was applied to the ULD series, the detection sensitivity increased for lesions ≥ 10 mm but it decreased for lesions 6 – 9 mm in size. With diffusion filtering, the detection sensitivity increased to 100% for lesions ≥ 10 mm, and it remained at 70% for lesions 6 – 9 mm in size. With diffusion, the detection sensitivity was not significantly different from that of the SD CTC

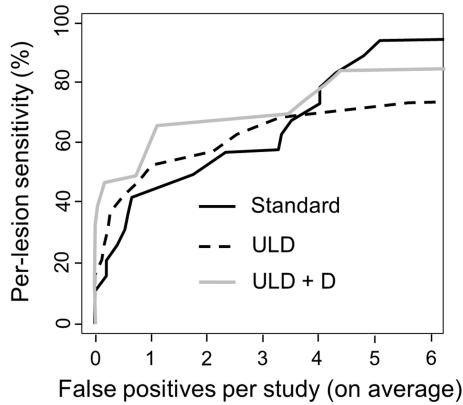


Fig. 6. Free-response receiver operating characteristic curve of the detection performance of CAD for polyps ≥ 6 mm in size on the testing set

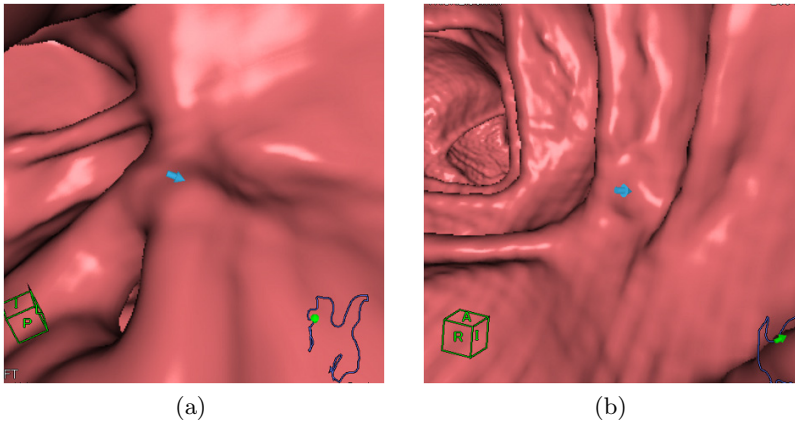


Fig. 7. Examples of missed lesions (indicated by blue arrows). (a) A 6-mm sessile polyp in colonoscopy that was missed with both standard-dose and ultra-low-dose CTC. (b) A 6-mm sessile lesion in colonoscopy that was missed with ultra-low-dose CTC but detected with standard dose CTC.

series. Fig. 6 depicts the detection accuracies in terms of free-response receiver operating characteristic curves.

4 Discussion

The results of this study indicate that the detection sensitivity of a CAD scheme that has been trained to detect polyps with conventional SD CTC studies can be reduced by 20% when it is used in ULD CTC. For large polyps, the detection

sensitivity can be improved by denoising of the image data. However, denoising by Gaussian filtering can further reduce the detection accuracy for polyps 6 – 9 mm in size. Therefore, the use of nonlinear diffusion filtering is preferred, because it can improve the detection sensitivity for large polyps without further degradation of the detection of smaller polyps.

According to the results, polyps that are seen clearly in SD CTC are likely to be detected also in ULD CTC. However, subtle polyps that are only 6 – 7 mm in their largest diameter, and/or that have flat morphology in CTC, may become more challenging to detect in ULD CTC than in SD CTC (Fig. 7). Although some previous studies have suggested that a low tube current does not have a significant effect on CAD performance, it is not clear if those studies considered the detection of challenging polyps. The effect of ULD CTC on CAD performance could also depend on the particular detection algorithm used by that CAD system.

Previous studies that considered the effect of ULD CTC on CAD involved phantoms or simulations. In this study, we used real clinical CTC studies. One of the limitations of the study is that the conspicuity of polyps could differ between the supine SD series and prone ULD series, thereby introducing a detection bias.

Another limitation is that we did not consider CTC studies with orally administered positive-contrast tagging. In recent years, orally administered tagging has become increasingly common in CTC, because it can be used to reveal polyps covered by fluid and to indicate feces that imitate lesions. However, if tagging has high attenuation in CTC images, it could increase image noise locally, thereby potentially impairing the detection of small and subtle polyps.

In addition to the reduction of tube current, novel CT reconstruction algorithms could be able to reduce the effective dose by >50% without significant effect on perceived image quality [7]. The radiation dose could also be reduced by careful monitoring of necessary CT examinations. A combination of these approaches is likely to provide the most optimal outcome for minimizing the radiation burden on patients.

Acknowledgements. This work was supported by grants from Prevent Cancer Foundation and NIH/NCI CA140934, CA095279, CA166816, and CA131718. We also thank Partners Research Computing for providing high-performance computing services.

References

1. Brenner, D., Georgsson, M.: Mass screening with CT colonography: should the radiation exposure be of concern? *Gastroenterology* 129, 328–337 (2005)
2. Chakraborty, D., Berbaum, K.: Observer studies involving detection and localization: modeling, analysis, and validation. *Med. Phys.* 31, 2313–2330 (2004)
3. Fisichella, V., Báth, M., Allansdotter, J., Jäderling, F., Bergsten, T., Persson, U., Mellingen, K., Hellström, M.: Evaluation of image quality and lesion perception by human readers on 3D CT colonography: comparison of standard and low radiation dose. *Eur. Radiol.* 20, 630–639 (2010)

4. Berrington de González, A., Kim, K., Landsdorp-Vogelaar, I., Rutter, C., et al.: Radiation-related cancer risks from CT colonography screening: a risk-benefit analysis. *Am. J. Roentgenol.* 196, 816–823 (2011)
5. Johnson, C., Chen, M.H., Toleday, A., Heiken, J., Dachman, A., Kuo, M., et al.: Accuracy of CT colonography for detection of large adenomas and cancers. *N. Engl. J. Med.* 359, 1207–1217 (2008)
6. Johnson, C., Dachman, A.: CT colonography: the next colon screening examination? *Radiology* 216, 331–341 (2000)
7. Kalra, M., Woisetschläger, M., Dahlström, N., Singh, S., Linblom, M., et al.: Radiation dose reduction with sinogram affirmed iterative reconstruction technique for abdominal computed tomography. *J. Comput. Assist. Tomogr.* 36, 339–346 (2012)
8. Kim, D., Pickhardt, P., Taylor, A., Leung, W., Winter, T., Hinshaw, J., Gopal, D., Reichelderfer, M., Hsu, R., Pfau, P.: CT colonography versus colonoscopy for the detection of advanced neoplasia. *N. Engl. J. Med.* 357, 1403–1412 (2007)
9. Kim, S., Lee, J., Shin, C., Kim, H., Lee, J., et al.: Effects of spatial resolution and tube current on computer-aided detection of polyps on CT colonographic images: phantom study. *Radiology* 248, 492–503 (2008)
10. Kiss, G., Drisis, S., Bielen, D., Maes, F., Van Cleynenbreugel, J., et al.: Computer-aided detection of colonic polyps using low-dose CT acquisitions. *Acad. Radiol.* 13, 1062–1071 (2006)
11. Laghi, A., Iafate, F., Rengo, M., Hassan, C.: Colorectal cancer screening: the role of CT colonography. *World J. Gastroenterol.* 16, 3987–3994 (2010)
12. Näppi, J., Frimmel, H., Dachman, A., Yoshida, H.: A new high-performance CAD scheme for the detection of polyps in CT colonography. In: Fitzpatrick, J., Sonka, M. (eds.) *SPIE Medical Imaging 2004*, vol. 5370, pp. 839–848 (2004)
13. Näppi, J., Yoshida, H.: Automated detection of polyps in CT colonography: evaluation of volumetric features for reduction of false positives. *Acad. Radiol.* 9, 386–397 (2002)
14. Näppi, J., Yoshida, H.: Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography. *Acad. Radiol.* 25, 287–300 (2007)
15. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE T. Pattern Anal.* 12, 629–639 (1990)
16. Pickhardt, P., Choi, J., Hwang, I., Butler, J., Puckett, M., Hildebrandt, H., Wong, R., Nugent, P., Mysliwiec, P., Schindler, W.: Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. *N. Engl. J. Med.* 349, 2191–2200 (2003)
17. Schoonenberg, G.A.F., Vries, A.H., de Grigorescu, S.E., Peters, J., Vilanova, A., Truyen, R., Stoker, J., Gerritsen, F.A.: Effects of filtering on colorectal polyp detection in ultra low dose CT. In: Manduca, A., Amini, A. (eds.) *SPIE Medical Imaging: Physiology, Function, and Structure from Medical Images*, vol. 6143, pp. 18–1–18–10. SPIE, San Diego (2006)
18. Thorne, M.: Background radiation: natural and man-made. *Radiol. Prot.* 23, 29–42 (2003)
19. Yoshida, H., Näppi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *IEEE Trans. Med. Imaging* 20, 1261–1274 (2001)

Application of CT Simulation Technique for Virtual Ultra-Low-Dose Trial in CT Colonography

Chang Won Kim^{1,4} and Jong Hyo Kim^{1,2,3,4,5}

¹ Interdisciplinary Program of Bioengineering, Seoul National University
College of Engineering, San 56-1, Sillim-dong, Gwanak-gu, Seoul 151-742, Korea
kcw11110@snu.ac.kr

² Department of Radiology, ³ Interdisciplinary Program in Radiation Applied Life Science,
Seoul National University College of Medicine

⁴ Institute of Radiation Medicine, Seoul National University Medical Research Center,
101 Daehangno, Jongno-gu, Seoul 110-744, Korea
kimjhyo@snu.ac.kr

⁵ Department of Intelligent Convergence Systems,
Graduate School of Convergence Science and Technology, Seoul National University,
864-1, Iui-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do 443-270, Korea

Abstract. A low-dose CT simulation technique is presented which might allow for a virtual ultra-low-dose trial in CT colonography without requiring raw sinogram data. A virtual sinogram is generated by performing the line integral of the CT number-based attenuation value with use of the CT scan parameters available in the DICOM header and in the literature. A separate noise sinogram is generated with use of a noise model, which incorporates the X-ray photon flux depending on the mAs, system electronic noise, and virtual sinogram. A synthetic noise CT image is generated by application of the filtered back projection of the noise sinogram with use of an appropriate filter that depends on the reconstruction kernel of the original CT. Finally, a simulated low-dose CT image is generated by addition of the CT data for the synthetic noise to the original CT data. Clinical CT colonography images with and without fecal tagging were used as simulation input and 50%, 25%, and 12.5% dose images were generated and evaluated. Our results suggest that the proposed CT simulation technique has potential for application in virtual ultra-low-dose trial in CT colonography in which an unlimited number of scan protocols could be performed without repetition of the real CT exposure to the patients.

Keywords: Algorithmic modulation transfer, noise power spectrum, low-dose simulation, filtered back projection.

1 Introduction

There is a growing concern about increased cancer risk due to the radiation exposure associated with CT examinations. This radiation-induced cancer risk is one of the limitations of CT colonography (CTC), which otherwise has a number of advantages over conventional methods for colon cancer screening and surveillance.

Therefore, various efforts are being made to reduce the amount of radiation in CT examinations while maintaining the diagnostic quality at a comparable level. Assessing the diagnostic performance of ultra-low-dose CTC is an example of studies, which explore the potential lower limit of low-dose CT imaging applications that do not compromise the diagnostic quality. However, for objective comparison, an ultra-low-dose study requires repeated CT scans of the same patients under different dose conditions, which involves ethical problems.

Low-dose CT simulation is a promising technique that allows the generation of simulated low-dose CT images that use patients' CT images without repeating the CT examinations to provide images at different dose levels. Although it was reported that low-dose CT simulation could be done by use of raw sinogram data provided by the CT vendor, it is very difficult to set up such an experimental setting in general academic institutions.

Our motivation was to develop a method which would enable low-dose CT simulation studies with use only of conventional CT image data and without the need for raw sinogram data, and then to evaluate its potential in CTC especially for virtual ultra-low-dose trial studies.

2 Materials and Methods

2.1 Materials

A 16-row CT scanner (Somatom Sensation 16, Siemens, Erlangen, Germany) was used in this study. Water-phantom and DCIOM images from the CTC study were used as the low-dose simulation input. We used the water phantom to validate the developed low-dose CT simulation technique. The phantom was scanned at 120 kVp and 200 mAs/30 mAs with eight different reconstruction kernels. The CTC images were scanned at 140 kVp with the current level around 40 mAs.

2.2 Methods

A low-dose CT simulation involves creating a noise image that reflects a set of CT parameters and the patient's anatomy. In this study, the required CT parameters were two sub-system functions, the algorithmic modulation transfer function (MTF) of the reconstruction kernels and the ramp filter apodization function; and two parametric values, the Q_0 (detector photon flux during the actual CT examination) and the system noise of the CT. The two sub-system functions were derived by use of the noise power spectrum (NPS) of the water phantom and its comparison to the ideal white Gaussian quantum noise. The two parametric values were obtained from the literature. Throughout the procedure, our assumption was that CT scans were performed in the axial mode, and reconstructions were carried out by use of the conventional filtered back projection (FBP) method.

2.2.1 Evaluation of the Sub-system Functions

Measurement of NPS. We made use of the NPS of CT as an information source regarding the system response of the CT in the spatial frequency domain. To measure the algorithmic MTF, we applied the subtracted NPS technique to the water phantom

scan dataset. It is known that the noise in CT images contains not only a quantum noise portion, but also a structural noise portion. Usually, the structural noise arises from scatter, dark current, non-uniform detector gain, beam hardening, shading, and other unknown factors. These types of noises are non-stochastic and can be classified as artifacts. Therefore, these structural noise portion can be canceled by subtraction of two images, which are scanned with the same phase. However, the resulting subtracted image should be divided by the square root of two, because subtracting the two images results in increasing the noise by the square root of two [3]. The NPS is give by

$$\text{NPS}_{2d}(u, v) = \frac{\Delta x \Delta y}{MN} \left| \text{FFT} \left\{ \frac{\Delta I(x, y)}{\sqrt{2}} \right\} \right|^2. \quad (1)$$

Derivation of the Algorithmic MTF. Usually, what we call MTF in CT means $\text{MTF}_{\text{total}}$, which can be divided by the algorithmic MTF (MTF_{alg}) and non-algorithmic MTF (MTF_{na}). Whereas the MTF_{na} is mainly determined by blurring due to the focal spot and aperture of the detector pixels [1], the MTF_{alg} solely reflects the frequency response of the CT system occurring in the FBP procedure. The $\text{MTF}_{\text{total}}$ can be written as

$$\text{MTF}_{\text{total}} = \text{MTF}_{\text{alg}} \text{MTF}_{\text{na}}. \quad (2)$$

From the fact that the NPS is driven only by the quantum noise, we can make use of the NPS as an information source to find a solution for MTF_{alg} . We calculated the MTF_{alg} for each reconstruction filter with equation (3), which describes the relationship among the NPS, ramp filter, and MTF_{alg} :

$$\text{NPS}(f) = \frac{\pi f}{\text{NEQ}} \text{MTF}_{\text{alg}}^2(f). \quad (3)$$

Ramp Filter Apodization Function. Because CT reconstructions are realized in a digital system, several sources of error are introduced during the digitization interpolation steps in the calculation of the reconstruction procedure. In order to prevent such errors, we evaluated the ramp filter apodization function, which reflects the decay of the ramp shape in the high-frequency range.

We created a sinogram consisting of white Gaussian noise, and we reconstructed it by using only the ramp filter to obtain a CT image that reflects the imperfection of the digital realization of the reconstruction. The NPS_{wgn} , which is obtained from this CT image, reveals the spectral response of the realization error. The ramp filter apodization function $H(f)$ is obtained by dividing the NPS_{wgn} by the ramp function shown in equation (4). Because this apodization occurs due to the digital realization process, the apodization function depends on the reconstruction pixel size and on the interpolation method used:

$$H(f) = \frac{NPS_{wgn}}{\text{ramp}(f)}. \quad (4)$$

Overall Reconstruction Function. Literally, FBP means that there is a filtering procedure before back projection, and therefore, the definition of the filter is a key factor. The overall reconstruction function is the composition of the ramp function, MTF_{alg} , and the ramp apodization function shown in equation (5).

$$|f| \cdot \frac{MTF_{alg-vender}(f)}{H(f)}. \quad (5)$$

2.2.2 Simulation of a Low-Dose CT Scan

The simulation of a low-dose scan requires the following steps:

- 1) A CT image in HU was converted to that of the attenuation coefficient by use of the effective μ_{water} at 140 kVp.
- 2) A virtual sinogram was produced through a virtual projection procedure, which performed the line integration of the attenuation coefficients along each ray path between the source and each detector pixel. As a result, the virtual sinogram $\{A_v(g, d)\}$ was created representing the linear attenuation at each gantry step and each detector location. The notation 'g' means each gantry step in CT, and 'd' denotes the location in the detector array.
- 3) The virtual sinogram data were translated to the virtual linear sinogram (S_v) by equation (6):

$$S_v(g, d) = Q_0 \cdot e^{-A_v(g, d)}. \quad (6)$$

- 4) The variance of the synthetic noise was calculated by equation (7). The first term on the right side is related to the quantum noise and the patient's anatomy, and the second term on the right side is related to the system noise.

$$\sigma_{sim}^2(g, d) = \alpha \cdot Q_0 \cdot T_v(g, d) \cdot \left(\frac{1}{\rho} - 1\right) + N_s. \quad (7)$$

In equation (7), α is a correction coefficient, which is calibrated by minimizing of the difference of the NPS curves between the original and the simulated image. Q_0 is the incident flux, which can be calculated with the TASMIP model. ρ is the ratio of mAS_{sim} to mAS_{org} .

$$N_s = b^2 \left(\frac{N_{0.Low}}{\rho^2} - N_{0.Hi} \right). \quad (8)$$

It has been reported that the system noise becomes a non-realizable factor when the linear attenuation is greater than 8. We referred the related equation for the CT scanner used in this study (Somatom Sensation 16, Siemens) to references [2, 5].

- 5) Synthetic noise was generated pixel by pixel by multiplying of the standard deviation with WGN, and was added to the virtual linear sinogram.

$$S_{\text{sim}}(\mathbf{g}, d) = S_v(\mathbf{g}, d) + \sigma_{\text{sim}}(\mathbf{g}, d) \cdot \text{WGN}. \quad (9)$$

WGN is the white Gaussian noise with '0' as the mean and the standard deviation of 1.

- 6) The noise-added virtual linear sinogram was then converted back to produce the virtual-noise sinogram, and was subtracted from the original virtual sinogram, producing the synthetic noise sinogram. This noise sinogram was filtered back projected and converted to HU to generate the synthetic noise CT image. Finally, the synthetic noise image was added to the original CT image, resulting in a simulated low-dose CT image.

3 Results

The results of the water phantom experiment are shown in Fig. 1. The similarity of the real low-dose CT (Fig. 1(b)) and simulated low-dose CT images (Fig. 1(d)) validates that the proposed simulation technique works properly both visually and quantitatively. The reconstructed noise image (Fig. 1(c)) with the proposed filter kernel confirmed that the proper noise pattern contained an appropriately mixed random-noise and streak-noise pattern.

Fig. 2 shows a comparison of the NPS graphs of the simulated and original CT images. Both the spectral pattern and the magnitude of the NPS graphs are comparable over the eight reconstruction kernels (B10f, B20f, B30f, and B40f).

Figs. 3 and 4 show the ultra-low-dose simulation results applied to the clinical CTC images. The original CTC cases were around 40 mAs, and the simulated ultra-low-dose level was around 50% (20 mAs), 25% (10 mAs), and 12.5% (5 mAs). Even though we could not compare the images to a real low-dose image, the simulated low-dose images were in good agreement upon visual comparison of the noise pattern with the ultra-low-dose scan images available from the literature. In Fig. 3, which is from a patient who had remaining tagging fluid, strong streak artifacts became conspicuous at ultra-low-dose levels of 10.5 mAs and 5.25 mAs, whereas in Fig. 4, which is from a patient who had no any remaining tagging fluid, no such streak artifacts were noticeable at the lowest level (5.875 mAs). This comparison reveals that the level of streak artifacts is associated with the amount of remaining tagging material in the CTC images. This finding suggests that the fecal tagging regimen should be considered when ultra-low-dose CTC studies are designed.

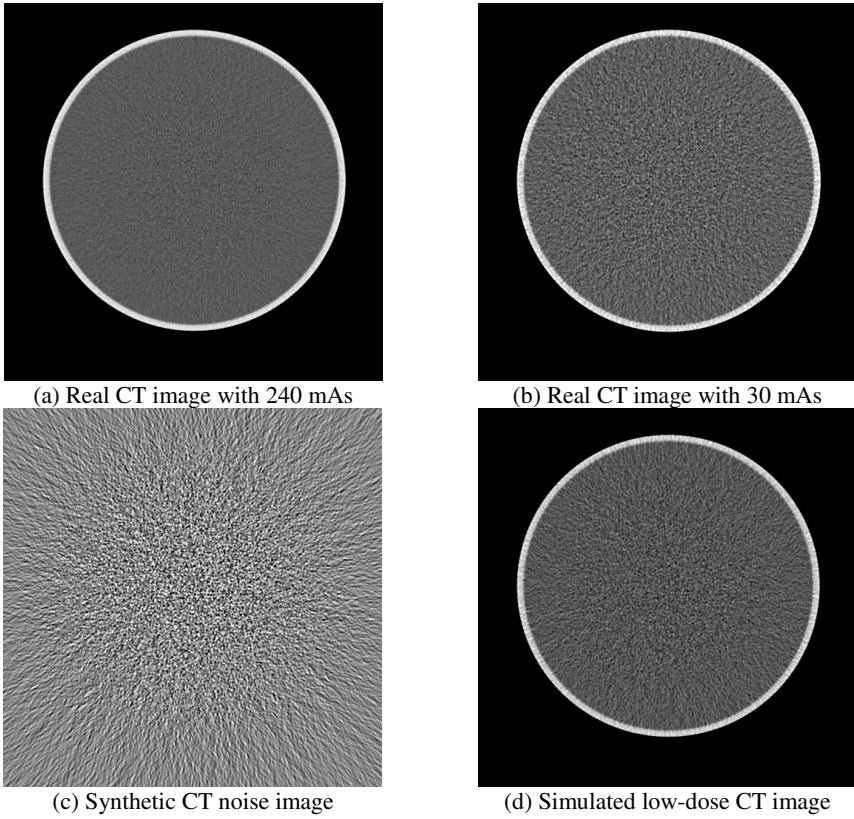


Fig. 1. Comparison of (a) high-dose scan (240 mAs), (b) low-dose scan (30 mAs), (c) simulated noise pattern, and (d) simulated low-dose (30 mAs) CT images for water phantom images

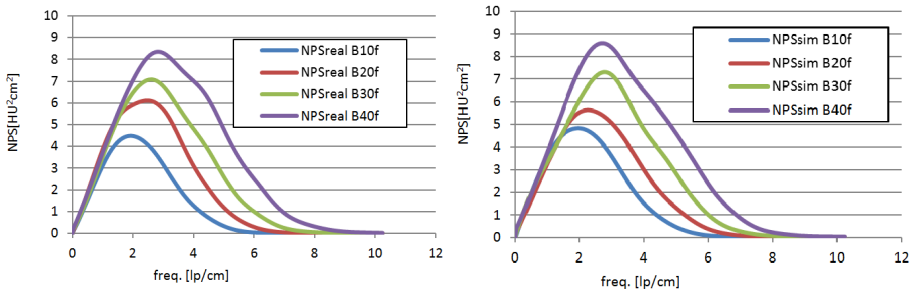


Fig. 2. NPS graphs of (a) real water phantom and (b) simulated low-dose CT image

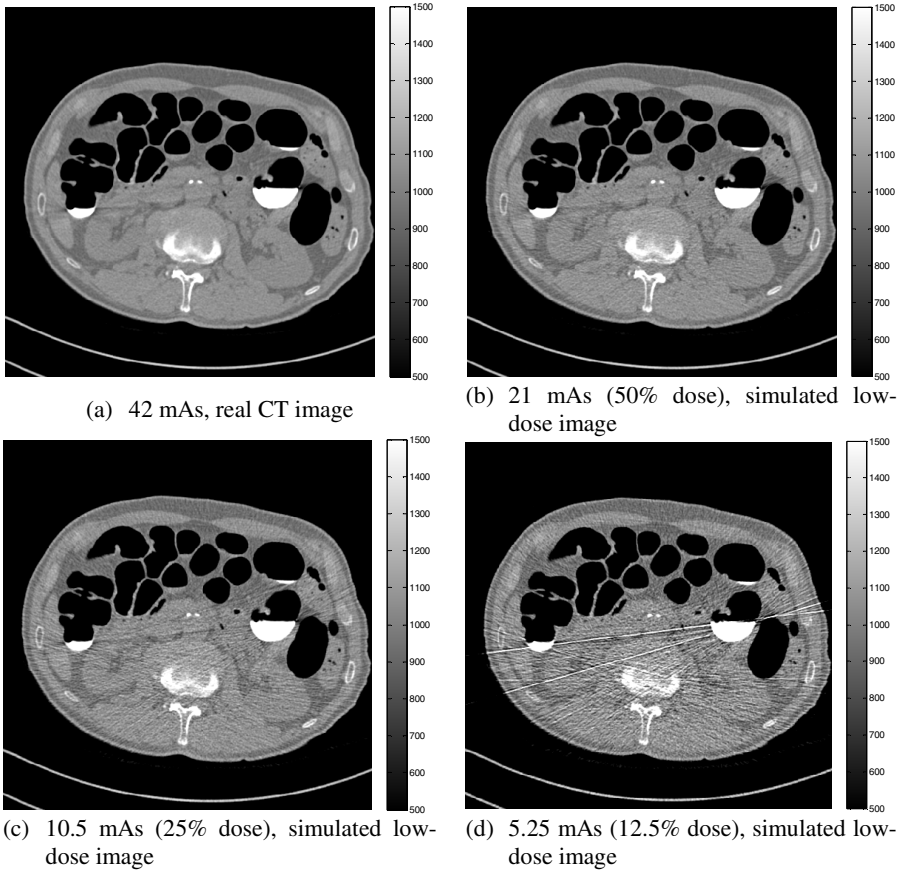


Fig. 3. Example of a resulting image by use of the low-dose simulation technique applied to a CTC case with tagging material remaining in the colon

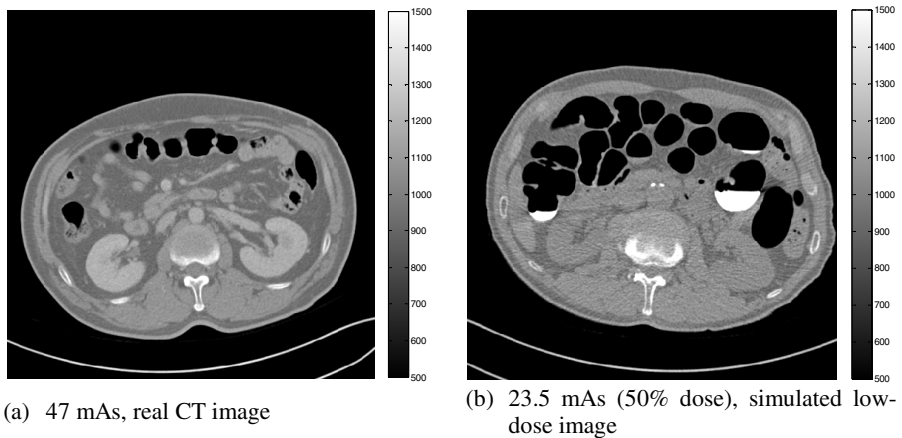


Fig. 4. Example of a resulting image by use of the low-dose simulation technique applied to a CTC case without any tagging material

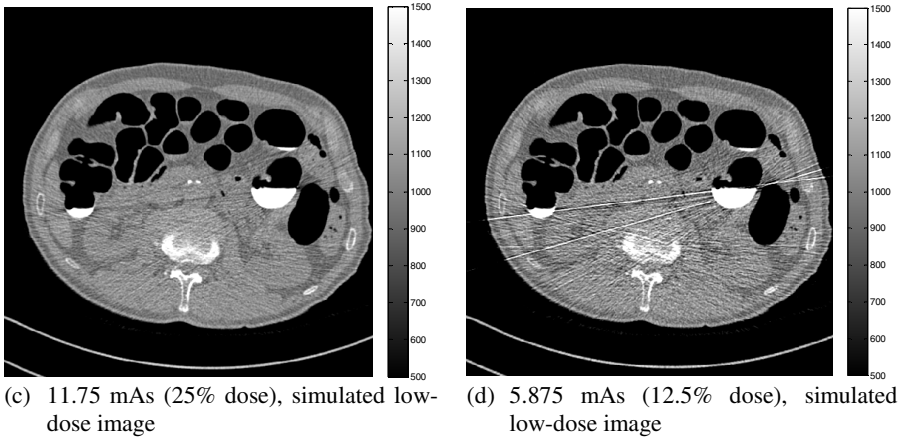


Fig. 4. (continued)

Table 1 compares the standard deviation (STD) measured at selected ROI areas including the kidneys and muscles marked on Fig. 5. Whereas the STDs of the 100% dose images were at a similar level, the STDs of the simulated 12.5% dose image for the patient with tagging material remaining in the colon were much higher than that of the patient without tagging material at the same dose level.

Table 1. Standard deviation of the ROI areas

Standard deviation					
Patient #1 (with tagging material)			Patient #2 (without tagging material)		
	42 mAs	5.25 mAs		47 mAs	5.875 mAs
ROI 1	26.5	185.6	ROI 4	20.6	99.5
ROI 2	23.0	193.2	ROI 5	23.4	94.4
ROI 3	18.3	109.9	ROI 6	19.2	66.8

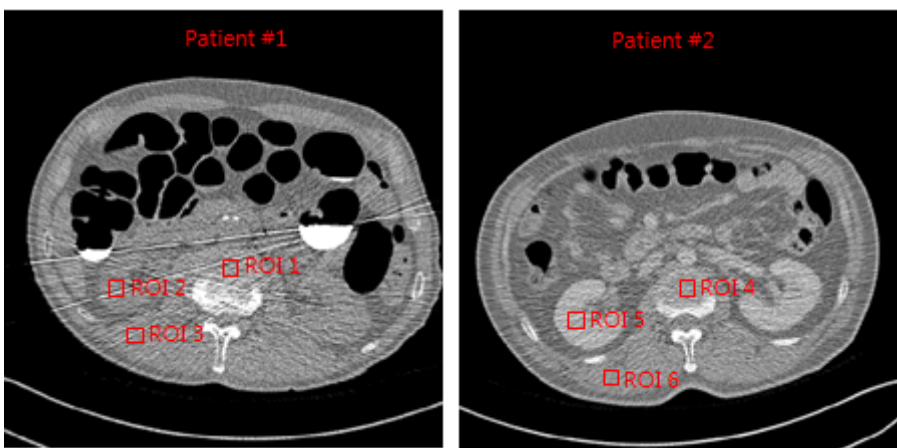


Fig. 5. ROI marks on the CT images of patients with (patient #1) and without (patient #2) tagging material

4 Conclusion

In this paper, we presented a low-dose CT simulation technique applied to both water-phantom and CTC images. The water phantom results verify that the pattern and magnitude of the generated noise on the simulated low-dose CT images were sufficiently similar to that of real low-dose CT scan images, and the simulated ultra-low-dose images for clinical CTC images provided a high level of realism in terms of noise and streak patterns.

One limitation of our study is that the tube current modulation was not included in our simulation procedure; this will be the subject of our future work.

Acknowledgements. This research was supported by the Converging Research Center Program through the Ministry of Education, Science and Technology (Grant No. 2011K000718)

References

1. Wagner, R.F., Brown, D.G., Pastel, M.S.: Application of Information Theory to the Assessment of Computed Tomography. *Med. Phys.* 6, 83–94 (1979)
2. Whiting, B.R., Massoumzadeh, P., Earl, O.A., O’Sullivan, J.A., Snyder, D.L., Williamson, J.F.: Properties of Preprocessed Sinogram Data in X-Ray Computed Tomography. *Med. Phys.* 33, 3290–3303 (2006)
3. Boedeker, K.L., Cooper, V.N., McNitt-Gray, M.F.: Application of the Noise Power Spectrum in Modern Diagnostic MDCT: Part I. Measurement of Noise Power Spectra and Noise Equivalent Quanta. *Phys. Med. Biol.* 52, 4027–4046 (2007)
4. Brenner, D.J., Hall, E.J.: Computed Tomography - An Increasing Source of Radiation Exposure. *N. Engl. J. Med.* 357, 2277–2284 (2007)
5. Massoumzadeh, P., Don, S., Hildebolt, C.F., Bae, K.T., Whiting, B.R.: Validation of CT Dose-Reduction Simulation. *Med. Phys.* 36, 174–189 (2009)

Application of CT Simulation Technique for Virtual Fecal Tagging in CTC

Zepa Yang¹, Hyeong-min Jin^{3,4}, and Jong Hyo Kim^{2,3,4,5}

¹ Department of Biomedical Sciences

² Department of Radiology

³ Interdisciplinary Program in Radiation Applied Life Science,
Seoul National University College of Medicine

⁴ Institute of Radiation Medicine, Seoul National University Medical Research Center,
101 Daehangno, Jongno-gu, Seoul 110-744, Korea
{yangzepa, hmjin, kimjhyo}@snu.ac.kr

⁵ Department of Intelligent Convergence Systems,
Graduate School of Convergence Science and Technology, Seoul National University,
864-1, Iui-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do 443-270, Korea

Abstract. An integrative CT simulation technique is presented that creates realistic CT images of virtual fecal-tagged material that was added to given clinical DICOM CT images. The energy spectrum of the CT X-ray source, the energy-dependent attenuation, and the scattering properties of the soft tissue and tagging material were incorporated in the generation technique for the DICOM image-based virtual sinograms, followed by CT reconstruction reflecting the vendor-specific filtering kernels. Dark band artifacts were generated by appropriate combining of beam-hardening and -scattering effects into the generation procedure for the virtual sinograms. We used a set of simple numerical phantoms to assess the basic behavior of artifact production. A reference set of CTC images with and without tagging material and artifacts was used for evaluation of the realism of the simulated results. The level of realism was evaluated in terms of the artifact strength and patterns around the added tagging material, compared to real tagging images. The results showed that our CT simulation technique provides sufficient realism for virtual fecal-tagged images that reflect a chain of physical and numerical processes, including beam hardening, scattering, and vendor-specific kernel filtered backprojection. The technique presented has the potential to be used as a tool for investigating the effect of tagging materials on image quality and to gauge how well the electronic cleansing technique performs.

Keywords: CT simulation, virtual tagging, CTC, artifact, beam hardening, scattering effect.

1 Introduction

CT colonography is a promising method in colon cancer screening, which provides a sensitivity comparable to that of optical colonoscopy for detecting polyps 10 mm and

larger [1]. Its ability to examine the whole-colon structure non-invasively in a time-efficient manner, without needing sedation and recovery time, places CTC as an attractive alternative to optical colonoscopy in the screening and surveillance steps of colon cancer diagnosis [1].

In an effort to reduce the patient's burden of the full cathartic bowel preparation currently required for CTC examination, studies are being carried out for the development of a noncathartic or a less intense cathartic bowel preparation in combination with an electronic bowel-cleansing technique [2]. In the research pathways towards optimizing the bowel preparation technique, a combination of studies would be needed before the final conclusion is reached, which would require huge costs and efforts, especially when varying image qualities with different tagging regimens are compared. Because tagging regimens with different materials and concentrations could generate varying degrees of artifacts in CT examinations, the performance of the visual examination as well as the electronic cleansing technique could be impaired depending on the resulting image quality due to the specific tagging regimen applied to the study. In addition, the image quality and the artifacts generated in CT are related to the patient's body structure in a complex manner; the same tagging regimen would generate different artifacts in different patients. Ideally, different tagging regimens should be applied to the same patients so that one can objectively compare the image quality and artifacts resulting from each specific tagging regimen; however, it is practically and ethically impossible.

The motivation for this study was to develop a technique that allows for virtual tagging on CTC images with different combinations of tagging materials and concentrations. This technique was applied to a CT simulation technique that included the energy-dependent attenuation characteristics of the tagging materials and of the body. We calculated the beam hardening and scattering process along the beam path within the body, which resulted in a realistic virtual tagging image with artifacts generated by the complex interactions of the tagging material with the X-ray energy and the body of the patient.

We describe the basic physical principle involved in beam hardening and scatter generation, present a simple experiment with numerical phantoms, and discuss the preliminary results in terms of clinical CTC cases.

2 Materials and Methods

2.1 Numerical Phantoms

A set of numerical phantoms that contained an oval-shaped disk of water and a circular disk of calcium was created, representing the human body and the spine, respectively, and multiple holes with air representing the insufflated colon. A hole was filled with iodine to varying levels of depths and concentrations to mimic the fecal tagging material. The numerical phantoms were used for assessment of the proper CT simulation technique.

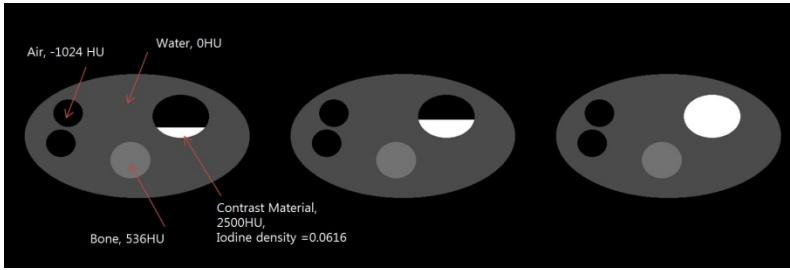


Fig. 1. Numerical phantom dataset. The intensities of each phantom are based on the parts of the CT images and contrast material and are based also on the iodine-based tagging regimen (OMNIPAQUE 300 mg/mL) and the density for the ratio in the clinical case. The window level of the image was changed to the best view of the HU value for each component.

2.2 Clinical Dataset

Selected CTC images were used from the CTC case database at the Seoul National University Hospital. CT examinations were performed with an MDCT (Siemens SOMATOM Definition, Philips Brilliance 64) under low-dose conditions (120 kVp, 40 ~ 55 mAs) and an iodine-based tagging regimen (OMNIPAQUE, 300mg/mL, Nycomed) was applied in the CTC cases used in this study.

We used DICOM images containing tagging material in the colon and exhibiting beam-hardening artifacts in the colon as a reference dataset to evaluate the degree of artifacts depending on the tagging concentration and the body configuration.

Fig. 2 shows sample reference images and line profiles revealing the degree of the dark-band artifacts around the tagging material. Table 1 summarizes the measured CT values at the tagging material and at a location with the most severe darkening artifact. As is shown, the degree of the artifacts has a complex relationship with the geometric configuration and concentration of the tagging material and with the body shape and the relative distance and location of another highly attenuating structure.

A separate set of DICOM images was collected which do not show tagging material or artifacts at a given slice, to be used as a substrate for addition of virtual-tagging material and for evaluation of the virtual-tagging effect.



Fig. 2. Various CT images with artifacts that lowered the tissue intensity

Table 1. Intensity of the contrast agent in the images with artifacts; the distorted intensity of the tissues compared to the normal tissues around the artifacts; the size of the contrast material; and the percent of contrast filling in the colon area. Distorted values are dependent on the contrast intensity of its source, its location, and the tissue material.

Intensity of the tagging material (HU)	Tissue intensity distortion (HU)	Length of the contrast material (mm)	Colon filling (%)
1186	-214	21 mm	13%
1787	-498	35 mm	33%
1579	-520	42 mm	21%
895	-115	24 mm	35%
1180	-211	16 mm	8%
1170	-282	27 mm	17%
972	-208	31 mm	30%
1050	-440	52 mm	25%

2.3 Simulation of X-Ray Interaction with Matter

2.3.1 CT Simulation

For a simple fan-beam CT simulation, an in-house CT projection/backprojection tool was created. The simulation software code was created with MFC (Microsoft Visual Studio was used). The basic algorithm for the projection/backprojection was based on the following equations:

$$P_{\gamma}(\xi) = \int_0^s \mu(\xi(x, y), \eta(x, y)) d\eta, \quad (1)$$

$$g(x, y) = \int_0^{\pi} P_{\gamma}(\xi) d\gamma, \quad (2)$$

where P_{γ} is the projection data (sinogram), μ is the original CT image, ξ the projection angle, η the count of projection rays, and g the backprojected image [3].

2.3.2 Beam Hardening Simulation

During the process of penetration through an object, the energy profile of a polychromatic X-ray beam changes continuously at each step of attenuation along the path. This is because different bands of the energy spectrum attenuate differently depending on the specific attenuation coefficient of the material being imaged. In general, low-energy radiation, such as soft X-ray beams, is more strongly absorbed than high-energy radiation, such as hard X-ray beams [3]. This beam-hardening effect is a major source of dark-band artifacts seen on CT images that contain contrast agents.

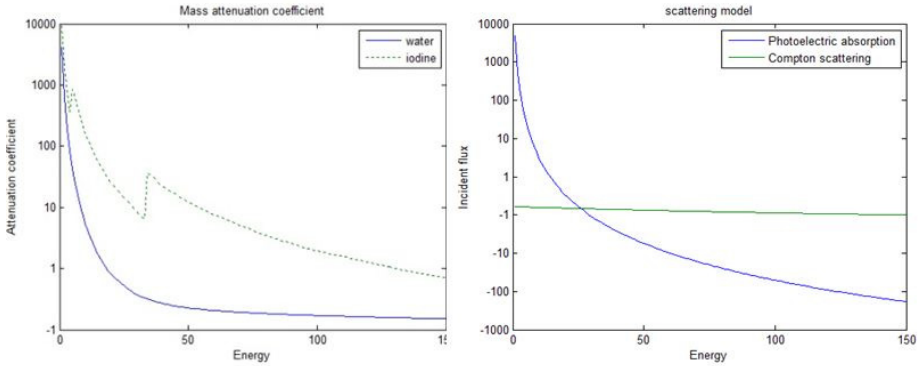


Fig. 3. A) Scattering model for total transmission, photoelectric absorption, and Compton scattering. B) Mass attenuation coefficients of water and iodine. These values were obtained from the National Institute of Standards and Technology (NIST) website.

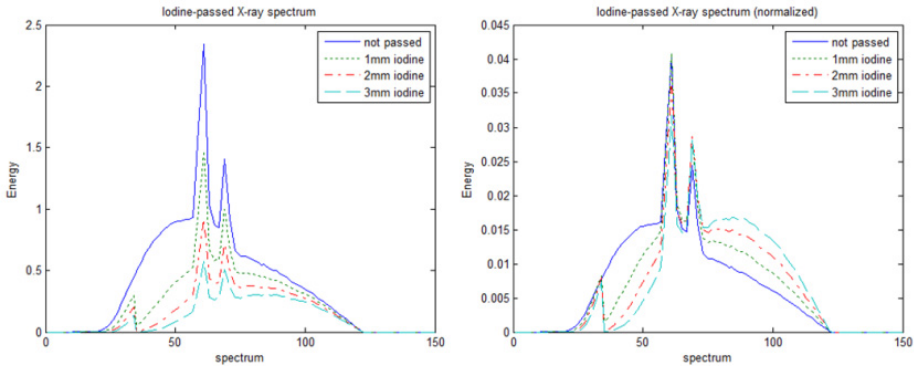


Fig. 4. A) Various thicknesses of iodine-passed X-ray characteristic spectrum and B) the normalized spectrum. A beam-hardening effect can be seen in the normalized spectrum; the low-energy part is degraded and the high-energy part is enhanced. The base X-ray characteristic spectrum was simulated with the tungsten anode spectral model by use of TASMIP, with 120 keV and a 3 mm aluminum filter.

Equation (3) describes the basic principle of X-ray attenuation used in this study for calculation of the beam-hardening process:

$$I_t(E) = \int I_0(E) e^{-\int \frac{\mu(E)}{\rho} x dx} dE, \quad (3)$$

where $\mu(E)$ is the energy-dependent mass attenuation coefficient, ρ is the density, and x is the thickness of the material that the x-ray spectrum has transmitted [3].

The X-ray spectrum was obtained from a published model for a tungsten anode spectral model by use of interpolating polynomials (TASMIP), with the application of 120 KVp and a 3 mm aluminum filter [4]. Beam hardening was calculated only for the iodine material, and the attenuation coefficient of water after passing the iodine material was updated by application of the hardened beam profiles to the spectral attenuation coefficients of water, as shown in Fig. 4.

X-ray attenuation is due to photoelectric absorption, Compton scattering, and Rayleigh scattering. The total spectral attenuation coefficient μ_{Total} is given by

$$\mu_{Total} = \mu_{photo} + \mu_{Compton} + \mu_{Rayleigh}, \quad (4)$$

$$\mu_{photo} = k \frac{\rho}{A} \frac{Z^4}{h\nu^3}, \quad (5)$$

$$\mu_{Compton} = n \cdot 2\pi r_e^2 \left[\left(\frac{1 + \epsilon}{\epsilon^2} \right) \left(2 \frac{1 + \epsilon}{1 + 2\epsilon^2} + \frac{\ln(1 + 2\epsilon)}{\epsilon} \right) + \frac{\ln(1 + 2\epsilon)}{2\epsilon} - \frac{1 + 3\epsilon}{(1 + 2\epsilon)^2} \right]. \quad (6)$$

For an analytical approximation, $\mu_{Rayleigh}$ was regarded as having a small, negligible contribution [5].

The values for the mass attenuation coefficients of iodine, water, aluminum, etc. were obtained from a publicly available database [7].

2.3.3 Scattering Simulation

Besides beam hardening, scattering is another important source of X-ray interaction with matter, which contributes strongly to the dark-band artifacts seen in CTC studies containing tagging materials. Although the Monte Carlo simulation technique is frequently used for calculation of the scattering process, it is impractically slow for use in CT simulation, which requires creating thousands of projection images per single CT image. Instead, we used a simple iterative-convolution model at each step of the ray path. The primary and the scattered transmission in water were calculated by use of the following equations:

$$T_p = \frac{I_t(x)}{I_o} = \int S(E) e^{-\int \mu_{photo}(total, E(\eta)) x d\eta} dE, \quad (7)$$

$$T_s = G \otimes \int S(E) e^{-\int \mu_{Compton}(water, E(\eta)) x d\eta} dE, \quad (8)$$

where G is the Gaussian filter convolution, x the thickness, and $S(E)$ is the weighting function derived by the following equation:

$$S(E) = \frac{I_o(E)}{\int I_o(E)}. \quad (9)$$

The relative fraction of the photoelectric absorption and the Compton scattering of water were used as weighting factors of the primary and scattered transmission for obtaining the total transmission at a given ray step:

$$T_t = T_p + k T_s, \quad (10)$$

where k is the scattering-factor constant.

After calculating the total transmission, we converted the value to a sinogram for the backprojection process.

Because the CT images are already filtered on the vendor-specific filtered backprojection process of their creation, if the images are projected and backprojected again, the images may have changes in the intensity and structure (i.e. blurring or over-sharpening by various filters). To avoid this, we acquired a sinogram from the original CT image, and we subtracted it from the sinogram of the virtual ray projection to yield a sinogram that only had the information for the beam-hardening effects and scattering artifacts. We backprojected the subtracted sinogram to acquire the image for the artifact part and finally added to the original CT image to get the artifact-simulated fecal-tagged image.

2.3.4 Integrated CT Simulation

To create realistic CT images containing tagging materials and artifacts, we applied an integrated CT simulation technique including image-based material decomposition, virtual sinogram projection, beam hardening and scatter simulation, and vendor-specific kernel filtering. Fig. 5 shows the overall procedure of our integrated CT simulation procedure.

In the image preparation step, the CT image of interest was used as a substrate; a tagging-component image was created by application of the attenuation coefficient and concentration of the tagging material; a water-component image was created by application of a threshold to the substrate image and adding the water part of the tagging-component image. Finally, a component sum image was created by addition of the tagging component image to the substrate image.

In the virtual-sinogram projection step, the ray projection was applied to the tagging component, water component, and component sub-images for calculation of the beam hardening, scattering, and photoelectric effects during ray penetration.

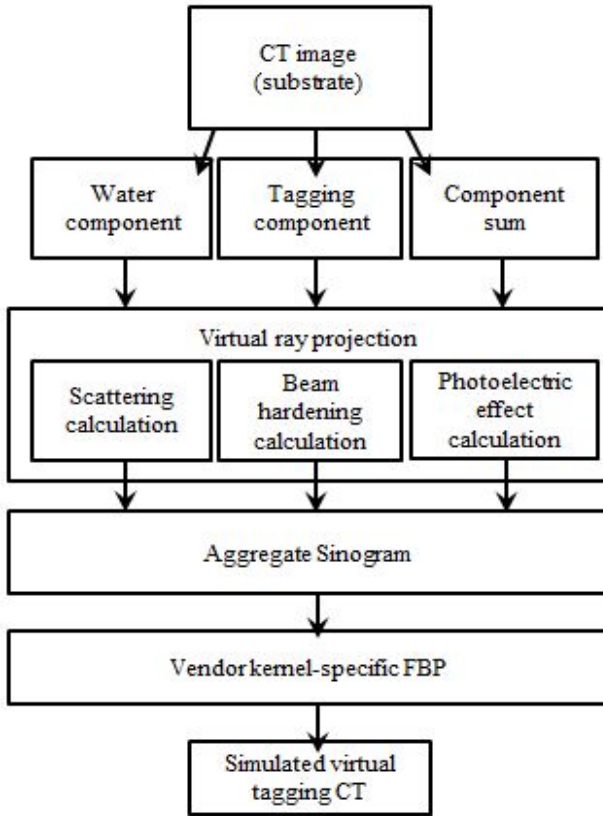


Fig. 5. Process box flowchart of the proposed method for simulating virtual fecal-tagged images

After the ray projection was completed, each component projection was weighted and added to create an aggregate sinogram, followed by vendor-specific kernel filtering, and finally backprojected to create the simulated CT of the virtual-tagging data.

3 Results

The initial simulation was tested with the numerical phantoms at various densities. Fig. 6 shows oval phantoms with various iodine fill levels and densities. The intensity and size of the artifacts varied depending on the density and the percent of iodine filling.

From the results of the evaluation, we found that the pattern of streak artifacts and their tissue intensity distortion were almost similar. In clinical trials, realistic CT images with artifacts were simulated successfully.

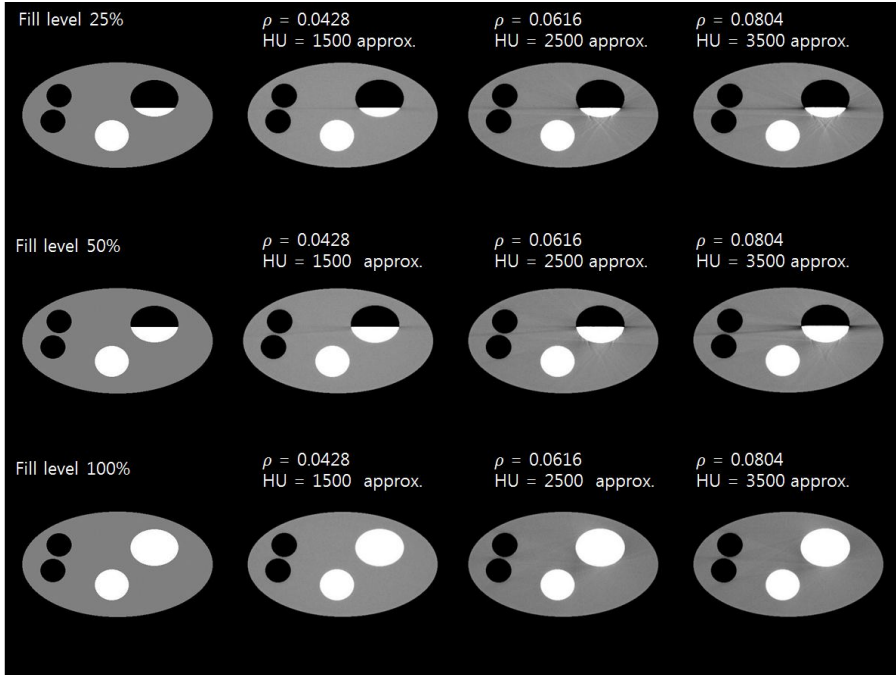


Fig. 6. Simulation results for the numerical phantoms with varying levels of colon filling and iodine density. The level of streak artifact varies depending on both iodine density and level of filling. The window width and the level of the image were adjusted to yield the best view of the artifacts.

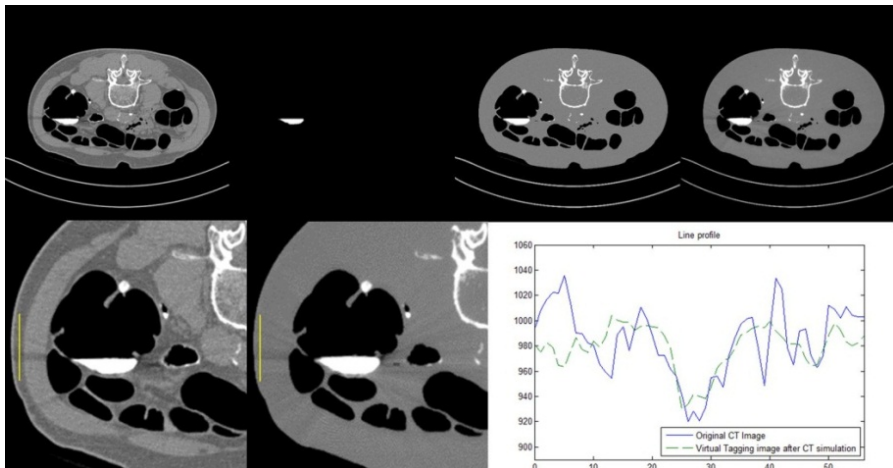


Fig. 7. A) CT image with artifacts, B) iodine mask, C) tissue normalized results with iodine mask, D) simulated result, and E) line profile of the original CT image and the virtual-tagging CT image result. Distorted values and their patterns are almost similar.

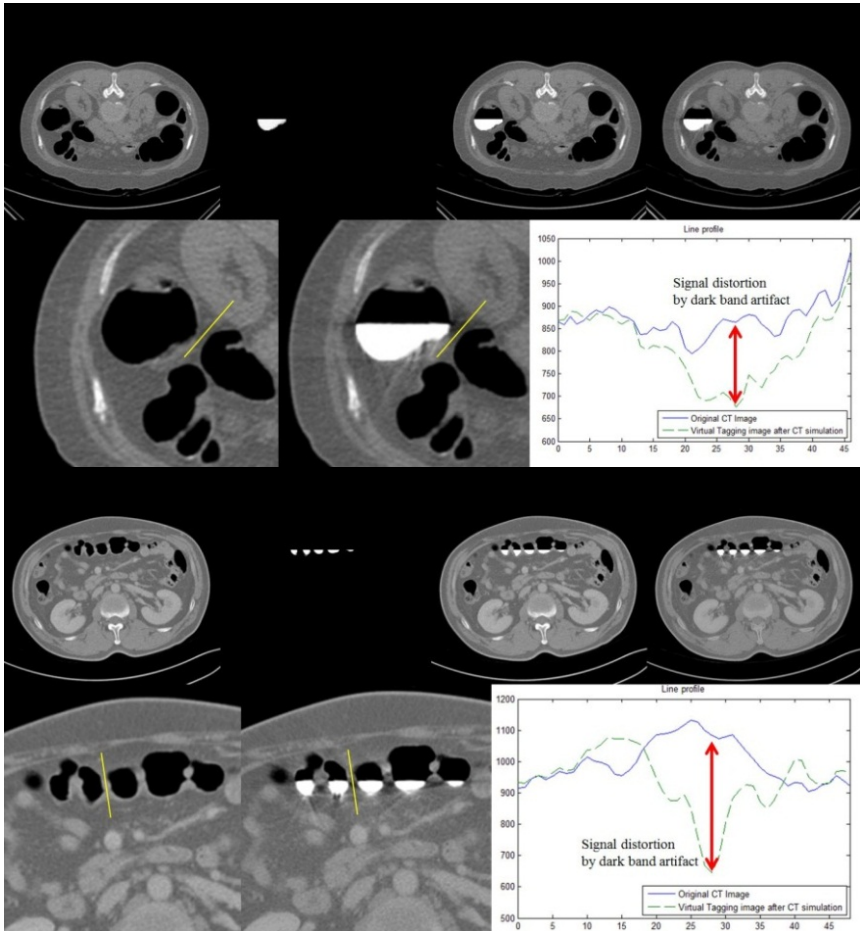


Fig. 8. A) Original CT images without artifacts, B) iodine-targeted image, C) iodine-mask added image, and D) CT simulated virtual fecal-tagged CT images. Beam-hardening and scattering artifacts were successfully generated. B) Line profile of the original CT image and virtual tagged image.

4 Conclusion

By use of the process described, virtual fecal-tagged CTC images can be generated at any location and for any circumstance. This result could contribute to the collection and practical use of patient data for research. Furthermore, it can be used as concrete datasets for CT colonography research in the future.

5 Discussion

This study presented an integrative CT simulation technique which could produce realistic dark-band artifacts caused by highly attenuating fecal tagging agents in CTC

studies. Dark-band artifacts are frequently seen in CTC studies with fecal tagging, and then often cause problems in electronic cleansing and polyp detection procedures in CAD systems. As shown in Table 1, the intensity of dark bands appears to be associated with several factors, including the density of the tagging agent, the level of fluid filling the colon, the width of filling, and the size of body. Although CT systems apply sophisticated techniques in the reconstruction procedure to reduce the dark-band artifacts, these artifacts still remain as a challenging problem not yet overcome.

Our study elucidated the physical process behind the generation of the dark-band artifacts. Whereas these artifacts are usually called beam-'hardening artifacts', our study showed that these artifacts are, in fact, generated by a combination of beam-hardening and scattering effects which are two fundamental processes of x-ray interaction with matter. In this study, we did not attempt to apply the proposed technique to removal of the dark-band artifacts. Although our simulation technique has such a potential, artifact removal would involve additional steps of specialized techniques which remain a subject for further study.

In its current form, however, the presented technique could be used in the evaluation of electronic cleansing techniques and polyp detection techniques in CTC under the fecal-tagging condition by providing an artifact-free image as the ground truth as well as the simulated virtual fecal-tagging image with dark-band artifacts as test data. This study is limited in that only fluid-state tagging material was considered. As various tagging regimens are being tried clinically, extending the tagging models of our simulation framework to include materials of solid or semi-solid forms which may appear as particles or thin surface layers would be necessary for advancement of this study to a more practical level.

References

1. McFarland, E.G., Levin, B., Lieberman, D.A., Pickhardt, P.J., Johnson, C.D., Glick, S.N., Brooks, D., Smith, R.A.: American Cancer Society, U.S. Multisociety Task Force on Colorectal Cancer, American College of Radiology: Revised Colorectal Screening Guidelines: Joint Effect of the American Cancer Society, U.S. Multisociety Task Force on Colorectal Cancer, and American College of Radiology. *Radiology* 248, 717–720 (2008)
2. Cai, W., Zalis, M.E., Näppi, J., Harris, G.J., Yoshida, H.: Structure-Analysis Method for Electronic Cleansing in Cathartic and Noncathartic CT Colonography. *Med. Phys.* 35, 3259–3277 (2008)
3. Buzug, T.M.: *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Springer, Heidelberg (2008)
4. Boone, J.M., Fewell, T.R., Jennings, R.J.: Molybdenum, Rhodium, and Tungsten Anode Spectral Models Using Interpolating Polynomials with Application to Mammography. *Med. Phys.* 24, 1863–1874 (1997)
5. Heismann, B.J., Leppert, J., Stierstorfer, K.: Density and Atomic Number Measurements with Spectral X-Ray Attenuation Method. *J. Appl. Phys.* 94, 2073–2079 (2003)
6. Ruhrschopf, E.P., Kligenbeck, K.: A General Framework and Review of Scatter Correction Methods in Cone Beam CT. Part 2: Scatter Estimation Approaches. *Med. Phys.* 38, 5186–5199 (2011)
7. Hubbell, J.H., Seltzer, S.M.: *Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients from 1 keV to 20 MeV for Elements Z=1 to 92 and 48 Additional Substances of Dosimetric Interest**, <http://www.nist.gov/pml/data/xraycoef/index.cfm>

Application of CT Acquisition Parameters as Features in Computer-Aided Detection for CT Colonography

Janne J. Näppi¹, Don Rockey², Daniele Regge³, and Hiroyuki Yoshida¹

¹ 3D Imaging Research, Department of Radiology,
Massachusetts General Hospital and Harvard Medical School,
25 New Chardon St., Suite 400C, Boston, Massachusetts 02114, USA
jnappi@partners.org, yoshida.hiro@mgh.harvard.edu

² University of Texas Southwestern Medical School,
5323 Harry Hines Blvd, Dallas, Texas 75390-9151, USA

³ Institute for Cancer Research and Treatment,
Candiolo Str. Prov. 142, IT-10060, Turin, Italy

Abstract. Studies have indicated that the acquisition parameters of computed tomography (CT) scans can have significant effect on the accuracy of computer-aided detection (CAD) in CT colonography. We investigated whether these parameters can be used as external features with conventional image-based features to improve CAD performance. A CAD scheme was trained with the CT colonography data of 886 patients, and it was tested with an independent set of 705 CT colonography cases. The results indicate that some CT acquisition parameters can be used successfully as features of the detected lesion candidates for improving the detection accuracy of CAD for flat lesions and carcinomas.

Keywords: Computed tomographic colonography, computer-aided detection, CT acquisition, polyp detection, virtual colonoscopy.

1 Introduction

Some of the acquisition parameters of computed tomography (CT) can have a significant effect on the detection accuracy of computer-aided detection (CAD) in CT colonography (CTC). Therefore, studies have been performed to determine the effect and optimal value of CT acquisition parameters for CAD in CTC. In two studies, the slice thickness of CTC images was observed to limit the size of the smallest polyp that is detectable by CAD [10,4]. In one study, CAD was applied to low-dose CTC by use of a 13 mAs/rotation, collimation of 1.5 mm, slice thickness of 3.0 mm, reconstruction interval of 1.5 mm, and table speed of 30 mm per rotation [3].

In practice, imaging devices are optimized for the purposes of visualization by radiologists rather than for the application of CAD. Therefore, it is likely that practical CAD systems will need to be able to adapt automatically to a wide variety of input data and parameter settings from multiple sources.

In a conventional application of CAD in CTC, the detected regions of lesion candidates are characterized in terms of shape-based and texture-based features. The features are collected into feature vectors that are analyzed by a classifier to determine the likelihood that a lesion candidate represents a true lesion. The detections with highest likelihood of being a lesion will represent the output of the CAD scheme.

In this study, we considered the application of CT acquisition parameters as additional external features of the lesion candidates. Because CT acquisition parameters affect the visual appearance of CTC images (Fig. 1), and because some of them are known to affect CAD performance anyway, we hypothesized that their use as additional features could improve the discrimination performance of the classifier used by a CAD scheme. By reflecting meaningful adjunct information about image contrast, volumetric isotropy, or smoothness of the reconstructed CTC image data, the CT acquisition parameters might be able to improve the efficacy of conventional image-based shape and texture features.

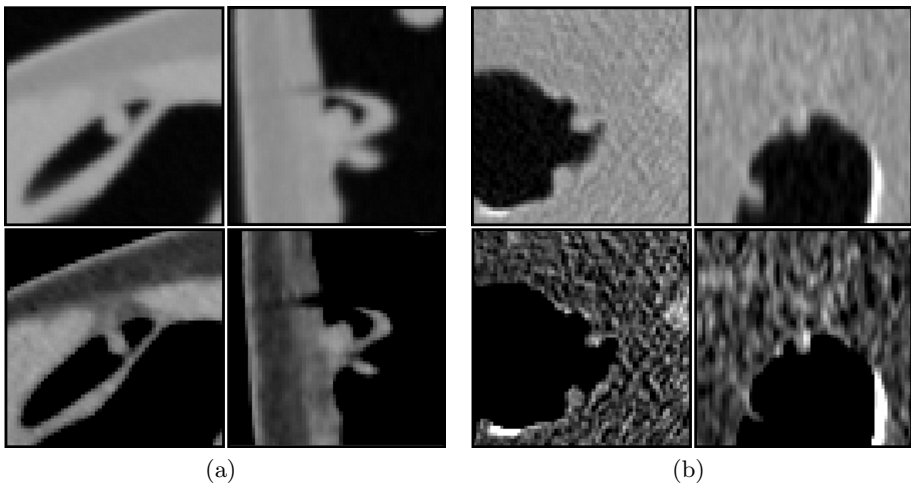


Fig. 1. Axial (left columns) and sagittal (right columns) CTC images of 7-mm polyps with wide (top rows) and narrow (bottom rows) display window settings. (a) Application of a soft reconstruction kernel, slice thickness of 1.0 mm, reconstruction interval of 0.8 mm, and tube current of 28 mA with 140 kVp. (b) Application of a soft reconstruction kernel, slice thickness of 2.5 mm, reconstruction interval of 1.25 mm, and tube current of 50 mA with 120 kVp.

The CT acquisition parameters that were used as external features in this study are described in Section 2.2. We considered only parameters that can be found in the file header of Digital Imaging and Communications in Medicine (DICOM) files. Furthermore, we considered only parameters that are likely to be associated with visually perceived image characteristics. For evaluation, each

external parameter was combined, one at a time, with a set of 6 conventional image-based shape and texture features that we have identified previously as the most effective and consistent features in polyp detection [7].

To provide a variety of examples of different CTC image acquisitions, and to avoid overfitting to a specific patient population or institution, we used a large number of CTC cases from multiple institutions. The evaluation was based on independent training and testing regimens. Statistical analysis was performed to assess the effect of the external features on the detection accuracy of CAD.

2 Methods

2.1 CAD Scheme

The fully automated CAD scheme that was used in this study processes an input CTC case in a number of steps. First, the CTC images are subjected to a pseudo-enhancement correction for minimizing tagging artifacts [8], and the volumetric data are interpolated to isotropic resolution [11]. Next, a thick region encompassing the colonic mucosa is extracted from input CTC data by use of a lumen-tracking method [7]. The locations of lesion candidates are determined by thresholding of lesion-like values of volumetric shape features of the extracted region [11]. The complete regions of detected lesion candidates are extracted by use of conditional morphological dilation [6]. Several shape and texture features are calculated for the regions of lesion candidates [5]. Finally, false-positive (FP) detections are reduced by use of a statistical classifier that determines the output of the CAD scheme [12].

Previously, we identified 6 shape and texture features that have emerged as consistently effective discriminative features over a wide variety of CTC populations [7]. In the following, we will denote this basic set of features as F6. For evaluation, each CT acquisition parameter (see Section 2.2) was combined with F6, one at a time, to yield sets of 7 features for the training and testing of the CAD scheme.

2.2 CT Acquisition Parameter Features

We considered the application of the following CT acquisition parameters as external features of lesion candidates:

- Slice thickness (ST). In a multi-detector array scanner, the ST is determined by binning of the different numbers of detector subunits together and by physically moving the collimator to the outer edges of the slices of a detector [1]. The image noise decreases when the slice thickness is increased.
- Reconstruction interval (RI). The RI defines the spacing of the reconstructed images. If the RI is too wide, small lesions that are located at the boundary between two slices could be missed due to partial-volume effects [9].
- Tube current (TC). A high TC reduces image noise, but it also increases radiation dose to the patient.

- In-plane image resolution (RS). The RS is the physical dimension of a pixel in a reconstructed CTC image.
- Focal spot (FS). The FS is the area on the anode of an x-ray tube that is struck by electrons and from which the resulting x-rays are emitted. An increment of focal spot size reduces the ability to define small structures.
- Exposure time (ET). The ET is the time during which the patient is exposed to radiation. A long scan time may cause breathing artifacts.
- Reconstruction kernel (K). The K that is convoluted with the image data determines the relationship between spatial resolution and image noise.
- Slice ratio (SR). The SR is a unitless feature that we derived from DICOM header information to characterize the isotropy of the reconstructed volumetric CTC data. It is calculated as $SR = RS/ST$. A small value of SR indicates that the volumetric image data may be suffering from geometric distortions, whereas a large value indicates that the physical voxels have nearly isotropic resolution, thereby minimizing distortions of image structures.

Most of these parameters have numerical values. However, the value of the reconstruction kernel (K) is highly dependent on the manufacturer and the model of the scanner. To provide a uniform value across different systems, we encoded K in terms of 3 values: 0 (soft reconstruction kernel), 1 (standard reconstruction kernel), and 2 (sharp reconstruction kernel).

To apply a CT acquisition parameter as an external feature, its value is provided as an additional feature in the feature vector of a lesion candidate (Fig. 2). It should be noted that all lesion candidates that are detected within the same CTC scan volume have the same value of a CT acquisition parameter feature. In this study, the external feature was combined with our usual set of conventional image-based shape and texture features (F6; see Section 2.1) of the lesion candidates to construct a vector of 7 features that is analyzed by the classifier of the CAD scheme to reduce FP detections and to determine the final output of the CAD scheme.

2.3 Materials

The training data included the CTC data of 886 patients from 17 institutions. The patients were prepared with cathartic bowel preparation. Orally administered positive-contrast tagging was used with 56 patients. The CTC acquisition was performed in supine and prone positions. There were 158 colonoscopy-confirmed carcinomas or adenomas (called hereafter as “advanced lesions”) in 112 patients: 56 were ≥ 10 mm and 102 were 6 – 9 mm in size.

The testing data included the CTC data of 705 patients from 13 institutions. The patients and institutions of the testing data were completely independent from those of the training data. The patients were prepared with cathartic bowel preparation. One third of the patients were administered positive-contrast tagging orally based on iodine alone or with barium. The CTC was performed in supine and prone positions. There were 260 colonoscopy-confirmed advanced lesions in 197 of the 705 patients: 158 of the lesions were ≥ 10 mm and 102 were 6 – 9 mm in size.

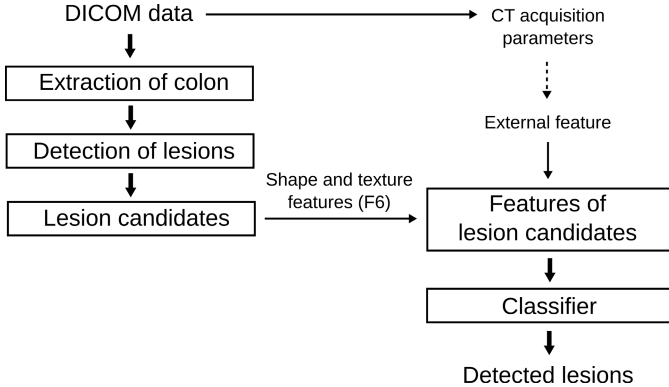


Fig. 2. Application of the CT acquisition parameters as external features within the CAD scheme. Each external feature was combined with the same 6 conventional image-based shape and texture features. A statistical classifier analyzes the features to reduce FP detections.

Table 1 summarizes the value range of the CT acquisition parameters between the training and testing cases.

Table 1. Minimum and maximum values of the CT acquisition parameters in the training and testing cases

Feature	Training		Testing	
	Min.	Max.	Min.	Max.
ST	1.0	5.0	1.0	5.0
RI	1.0	5.0	0.59	2.0
TC	50	408	28	300
RS	0.50	0.97	0.51	0.97
FS	0.7	1.2	0.7	1.2
ET	27	3100	478	1825
K	0	1	0	1
SR	0.10	0.72	0.14	0.79

2.4 Evaluation

To provide clinically meaningful results, we limited the number of detections that can be displayed by the CAD scheme to a maximum of 15 per patient. It is likely that a display of a larger amount of CAD detections, most of which are FP detections, would make the use of computer-assisted interpretation too tedious for clinical practice. Given this constraint, we determined the maximum detection sensitivity and the median number of FP detections per patient.

To assess the statistical significance of the effect of a feature combination on detection accuracy, a pair-wise randomization test was performed by comparing the detection accuracies of CAD using the conventional 6 features (F6) with that of using the F6 together with one of the CT acquisition parameter features. The figure-of-merit (FOM) was the partial area under the free-response receiver-operating characteristic (FROC) curve, where the partial area was calculated to the left from the point where either CAD scheme reached its maximum sensitivity [2].

3 Results

Table 2 shows the per-lesion detection accuracy of CAD in the testing set for all advanced lesions ≥ 6 mm. The first row (F6) shows the result with the 6 conventional features, whereas the other rows show the result of an indicated combination of 7 features. The use of the SR-feature yielded highest increment in detection sensitivity, but the improvement was not statistically significant.

Table 2. Per-lesion detection accuracy of CAD for advanced polyps and flat lesions. Arrows indicate improvement of detection sensitivity as compared with the original feature set (F6; first row).

Features	Sensitivity FPs/case		Sensitivity FPs/case	
	≥ 10 mm	median	6 – 9 mm	median
F6	91%	11	81%	12
F6 + SR	↑92%	15	↑85%	15
F6 + TC	91%	14	77%	15
F6 + ST	↑92%	13	↑84%	12
F6 + RI	↑92%	12	79%	↓9
F6 + K	↑92%	11	↑83%	↓11
F6 + FS	↑92%	↓10	81%	↓10
F6 + RS	91%	12	80%	12
F6 + ET	↑92%	↓10	80%	13

For polyps ≥ 6 mm, the small improvement in detection accuracy by use of the SR, from 0.71 [0.706, 0.714] to 0.71 [0.710, 0.717] (the numbers in brackets indicate 95% confidence intervals), was not statistically significant. However, for flat lesions ≥ 6 mm, the use of SR improved the overall per-lesion detection accuracy significantly from 0.68 [0.674, 0.690] to 0.70 [0.689, 0.702] ($p < 0.05$).

Table 3 shows the detection result for advanced flat lesions. Again, the use of the SR yielded most improvement in detection sensitivity. For flat lesions 6 – 9 mm, the improvement in overall detection accuracy from 0.61 [0.595, 0.624] to 0.64 [0.628, 0.654] was statistically significant ($p < 0.01$).

Table 3. Per-lesion detection accuracy for advanced flat lesions. Arrows indicate improvement of detection performance as compared with the original feature set (F6; first row).

Features	Sensitivity	FPs/case	Sensitivity	FPs/case
	≥ 10 mm	median	6 – 9 mm	median
F6	100%	6	75%	5
F6 + SR	100%	↓4	↑81%	7
F6 + CR	100%	6	63%	14
F6 + ST	100%	↓4	75%	6
F6 + RI	100%	↓4	75%	7
F6 + K	100%	10	75%	11
F6 + FS	100%	8	75%	6
F6 + RS	100%	↓4	75%	8
F6 + ET	100%	8	75%	7

The improvement of overall per-lesion detection accuracy by the SR-feature was highly significant for carcinomas ≥ 6 mm. The detection accuracy improved from 0.77 [0.759, 0.773] to 0.83 [0.827, 0.837] ($p < 0.000001$). In contrast, for adenomas, the overall per-lesion detection accuracy was reduced marginally but significantly, from 0.77 [0.771, 0.777] to 0.77 [0.762, 0.768] ($p < 0.001$).

4 Discussion

The results indicate that the acquisition parameters of CTC scans can indeed be used successfully as external features to improve the detection accuracy of CAD. However, only few parameters yielded a meaningful improvement, and the precise effect depends on the type of target lesion. In particular, only the SR-feature yielded statistically significant improvements. With SR, the detection accuracy improved significantly in the detection of flat lesions ≥ 6 mm in size, advanced flat lesions 6 – 9 mm in size, and carcinomas ≥ 6 mm in size.

It is not particularly surprising that the use of the SR-feature yielded meaningful improvement in CAD performance. In previous studies, CAD performance has been found to be highly dependent on z-axis spatial resolution. The SR features appears to provide meaningful adjunct information about the anisotropy of physical image resolution, thereby improving CAD performance [4].

We used CT acquisition parameters that can be acquired from the header data of DICOM files. A limitation of this approach is the confusion and inconsistency of the terminology that is used by the different manufacturers and their scanner models. For example, some operating modes of the scanners can yield different effective values than what is specified by the DICOM header, and a parameter value may have different effects on different scanners. Therefore, for a practical application, a more detailed scanner-specific analysis of the optimal parameter values may be needed.

Another limitation of the study is that we did not consider combinations of CT acquisition parameter features but we analyzed the effect of each feature one at a time. In combination, they could have a greater impact on improving the detection accuracy of CAD.

5 Conclusion

We investigated the application of CT acquisition parameters as external features with conventional image-based features for improving the detection accuracy of CAD in CTC. A CAD scheme was trained and tested with large independent sets of clinical CTC cases from multiple institutions. The results indicate that the use of the so-called slice-ratio feature in particular can yield significant improvement in the detection of flat lesions and carcinomas.

Acknowledgements. This work was supported in part by grants of Prevent Cancer Foundation, NIH R21CA140934, CA095279, CA166816, and CA131718. We thank Dr. Richard Choi (Virtual Colonoscopy Center, Walter Reed Army Medical Center, Washington, DC, USA) for providing CTC cases for this study. We also thank Partners Research Computing for providing high-performance computing services.

References

1. Bushberg, J., Seibert, J., Leidholdt Jr., E., Boone, J.: The essential physics of medical imaging, 2nd edn. Lippincott Williams & Wilkins, Philadelphia (2002)
2. Chakraborty, D.: Analysis of location specific observer performance data: validated extension of the jackknife free-response (JAFROC) method. *Acad. Radiol.* 13, 1187–1193 (2006)
3. Chowdhury, T., Whelan, P., Ghita, O., Sezille, N., Foley, S.: Development of a synthetic phantom for the selection of optimal scanning parameters in CAD-CT colonography. *Med. Eng. Phys.* 29, 858–867 (2007)
4. Kim, S., Lee, J., Shin, C., Kim, H., Lee, J., et al.: Effects of spatial resolution and tube current on computer-aided detection of polyps on CT colonographic images: phantom study. *Radiology* 248, 492–503 (2008)
5. Näppi, J., Yoshida, H.: Automated detection of polyps in CT colonography: evaluation of volumetric features for reduction of false positives. *Acad. Radiol.* 9, 386–397 (2002)

6. Näppi, J., Yoshida, H.: Feature-guided analysis for reduction of false positives in CAD of polyps for CT colonography. *Med. Phys.* 30, 1592–1601 (2003)
7. Näppi, J., Yoshida, H.: Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography. *Acad. Radiol.* 25, 287–300 (2007)
8. Näppi, J., Yoshida, H.: Adaptive correction of the pseudo-enhancement of CT attenuation for fecal-tagging CT colonography. *Med. Image Anal.* 12, 413–426 (2008)
9. Prokop, M., Galanski, M.: *Spiral and multislice computed tomography of the body.* Thieme Verlag, Ludwigsburg (2003)
10. Sundaram, P., Beaulieu, C., Paik, D., Schraedley-Desmond, P., Napel, S.: CT colonography: does improved z resolution help computer-aided polyp detection? *Med. Phys.* 30, 2663–2674 (2003)
11. Yoshida, H., Näppi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *IEEE Trans. Med. Imaging* 20, 1261–1274 (2001)
12. Yoshida, H., Näppi, J., MacEaney, P., Rubin, D., Dachman, A.: Computer-aided diagnosis scheme for detection of polyps at CT colonography. *Radiographics* 22, 963–979 (2002)

Comparative Performance of State-of-the-Art Classifiers in Computer-Aided Detection for CT Colonography

Sang Ho Lee, Janne J. Näppi, and Hiroyuki Yoshida

3D Imaging Research, Department of Radiology,
Massachusetts General Hospital and Harvard Medical School,
25 New Chardon St., Suite 400C, Boston, Massachusetts 02114, USA
{lee.sangho,yoshida.hiro}@mgh.harvard.edu, jnappi@partners.org

Abstract. Several effective machine learning and pattern recognition schemes have been developed for medical imaging. Although many classifiers have been used with computer-aided detection (CAD) for computed tomographic colonography (CTC), little is known about their relative performance. This pilot study compares the performance of several state-of-the-art classifiers and feature selection methods in the classification of lesion candidates detected by CAD in CTC. There were four classifiers: linear discriminant analysis (LDA), radial basis function support vector machine (RBF-SVM), random forests (RF), and gradient boosting machine (GBM). There were five feature selection methods: sequential forward inclusion (SFI) of principal components (PCs), univariate filtering (UF), UF of PCs, recursive feature elimination (RFE), and RFE of PCs. A strategy of using all available features was tested also. For evaluation, 232,211 detections by a CAD system on 1,211 patients were subsampled randomly to create 10 different populations of 500 true-positive (TP) and 500 false-positive (FP) detections. The classifier performance was evaluated by use of the area under the receiver operating characteristic curve of 3 repeated 10-fold cross-validations. According to the result, the discrimination performance of the RBF-SVM classifier with feature selection by the RFE of PCs compared favorably with other methods, although no single classifier outperformed other classifiers under all conditions and feature selection schemes.

Keywords: Classification, feature selection, comparative performance, machine learning, virtual colonoscopy.

1 Introduction

Computed tomographic colonography (CTC) is a promising alternative to traditional invasive colonoscopy methods used in the detection and removal of polyps of the colon [1-3]. Computer-aided detection (CAD) systems for CTC typically make use of a classifier to discriminate between true-positive (TP) and false-positive (FP) findings generated by a polyp candidate detection system based on a set of features extracted

from the candidates [4-6]. However, CAD systems for CTC still display large numbers of FP detections [7]. Consequently, improving the detection specificity of CAD remains a challenging task in CTC, and a powerful classification engine is needed to deal with this difficult classification problem [8, 9].

The aim of a classification system is to classify an input pattern presented to the system to a correct category based on a feature vector of the input pattern. The complexity of the classification problem relies on the variability of the feature values for patterns in the same class relative to the difference between feature values for patterns in different classes. As a result, the optimality of a classifier depends on a specific dataset [10]. Thus, the goal of achieving the optimal performance for a pattern recognition system may be inconsistent with obtaining the best performance for a single classifier, which may also be associated with different feature selection schemes. This pilot study compared several state-of-the-art classifiers and feature selection schemes by using a large database in the classification task for CAD in CTC.

2 Method

2.1 Feature Selection

The goal of feature selection is to select a subset of relevant features for building robust classifiers by removing irrelevant and redundant features from input data. This is expected to improve the speed of construction and the accuracy of the final classifier.

From a theoretical perspective, it can be shown that optimal feature selection for supervised learning problems requires an exhaustive search of all possible subsets of features. However, for a large number of features or samples, an exhaustive search for an optimal feature set is impractical. Therefore, instead of an optimal set, in practice a supervised learning algorithm searches for a satisfactory approximation of the optimal set of features for a particular classifier.

In this study, three principal state-of-the-art methods were considered for feature selection, including 1) principal component analysis (PCA) [11], 2) univariate filtering (UF) [12], and 3) recursive feature elimination (RFE) [13].

Principal Component Analysis. The PCA is a well-established method for feature extraction and dimensionality reduction. It is based on the assumption that most information about features is contained in the directions along which the variation of the features is largest. The most common derivation of PCA is a standardized linear projection, which maximizes the variance in the projected space.

Univariate Filtering. UF is a feature selection method that reviews the features by using univariate statistical methods, such as the *t*-test or ANOVA models, to assess the efficacy of each individual feature in class prediction. UF is relatively dominantly used because of its simplicity and efficiency. However, it does not take into account feature-feature interactions, possibly leading to less accurate classifiers. UF is based on including the highest-ranked individual features depending on a chosen association

measure. Since UF applies independent evaluation criteria without the process of discovering patterns in data, it does not inherit any bias of a learning algorithm and it is also computationally efficient. UF is preferred in applications where application of data mining algorithms would be too costly or unnecessary in dealing with high dimensional features.

Recursive Feature Elimination. RFE is a multivariate approach based on the information content of a group of features, which uses successive elimination of individual features ranked lowest according to a criterion, aimed at keeping the discrimination ability as high as possible. It attaches a weight to each available feature. Based on the assumption that the features with the smallest weights are least informative in a feature set, a predefined number of features is removed iteratively from the set of available features. RFE involves combinatorial searches through the space of feature subsets, guided by the prediction ability of a specific classification model. Since grouping and predictive analysis of multidimensional features are used to control the selection of feature subsets, RFE tends to give superior performance as feature subsets found are better suited to the predetermined learning algorithm. Consequently, it is more computationally expensive than the UF.

In this study, a total of five feature selection methods derived from PCA, UF, and RFE were considered: 1) sequential forward inclusion (SFI) of the principal components (PCs) of PCA, 2) UF, 3) UF of PCs, 4) RFE, and 5) RFE of PCs. For comparison, also a strategy of using all available features without explicit feature selection was considered.

2.2 Classification

The goal of classification is to identify the correct category of an input pattern. The classification is typically based on an initial training set of samples whose category is known.

In this study, the following state-of-the-art classifiers were considered: 1) linear discriminant analysis (LDA) [11], 2) radial basis function (RBF) support vector machine (RBF-SVM) [14], 3) random forest (RF) [15, 16], and 4) gradient boosting machine (GBM) [17, 18]. Each classifier, except for LDA, evaluated the effect of its model tuning parameters by using resampling. Optimal tuning parameters were chosen across those parameters. Finally, the classification performance was estimated from a training set.

Linear Discriminant Analysis. LDA is a robust and fundamental classifier. It is used for finding an optimal transformation that maps input data into a lower dimensional space to minimize the within-class distance and simultaneously to maximize the between-class distance, thus achieving maximum discrimination. LDA is closely related to PCA in that both look for linear combinations of features which best explain the data. LDA attempts to model the difference between classes of data explicitly, whereas the PCA does not consider differences between classes.

Support Vector Machine. The SVM is based on the concept of decision planes that define boundaries. A decision plane is one that separates between a set of features having different class memberships. The classification is based on separating hyperplanes that distinguish between objects of different class memberships in a multi-dimensional space. The basic idea behind the SVM is to create nonlinear boundaries by generating linear boundaries on a higher-dimensional space, where the original features are rearranged by use of a set of mathematical functions known as kernels. There are a number of kernels that can be used in SVM models, including linear, polynomial, sigmoid, and RBF kernels. The RBF-SVM is the most popular choice among the kernel types used in the SVM.

Random Forest. The RF classifier is an ensemble of decision trees, which combines the predictions of many classification trees to obtain more accurate classifications. Many samples of the same size as the original dataset, called bootstrap samples, are drawn from the dataset with replacement. In each bootstrap sample, approximately 68% of the observations in the original dataset occur one or more times. The observations in the original dataset that do not occur in the bootstrap sample are said to be out-of-bag for that bootstrap sample. For each bootstrap sample, a decision tree is built. At each step of the building process, only a small number of variables are available for construction of the decision tree. There is no pruning of the decision trees of a RF classifier. The trees of the RF are then used for constructing predictions for all out-of-bag observations of bootstrap samples. The predicted class of an input sample is acquired by voting for the predicted class among all the trees.

Gradient Boosting Machine. Boosting is a process that combines many separate prediction rules, some of which may be quite weak on their own, to produce a more powerful combined classifier. The GBM is another procedure that, like the RF, fits many trees to a single dataset. The GBM differs from the RF in that the trees are built sequentially, with observation weights updated according to whether the observations are correctly or incorrectly classified. Boosting iteratively adds basis functions in a greedy fashion such that each additional basis function further reduces the selected loss function. The GBM is one of the more novel classifiers that, to date, has rarely been applied in the analysis of medical images.

2.3 Materials and Evaluation

The empirical data for this study included potential lesion candidates detected by a CAD system [19] from a large clinical CTC screening population of 1,211 patients at 20 medical centers [20]. The patients were prepared cathartically for the CTC examination. Orally administered fecal tagging was used for 37% of the patients. The CTC data were acquired by use of 11 CT scanners with an average slice thickness of 2.35 mm (range, 1.0 – 5.0 mm) and average current of 156 mA (range, 50 – 408 mA). Approximately 18% of the patients had clinically significant colonoscopy-confirmed lesions. There were 317 lesions ≥ 6 mm: 40% of the lesions measured ≥ 10 mm and

60% measured 6 – 9 mm in the largest diameter. Approximately 84% of the lesions had polypoid morphology and 16% had flat morphology.

There were 232,211 CAD detections, including 929 TP detections and 231,282 FP detections. Because some of the lesions were detected multiple times, the number of CAD detections is higher than that of confirmed true lesions in the patients. The detections were sampled randomly without replacement for construction of 10 population samples for an unbiased evaluation of classifier performance under various conditions. Each subsample contained 500 TP and 500 FP CAD detections. Each detection was characterized by a total of 67 shape and texture features.

The classifier performance was evaluated by use of three repeated 10-fold cross-validations, where the performance was measured by use of the area under the receiver operating characteristic curve (A_z). The A_z was assessed for each population sample by use of the four different classifiers with each of the five different feature selection schemes. Fig. 1 illustrates the study design.

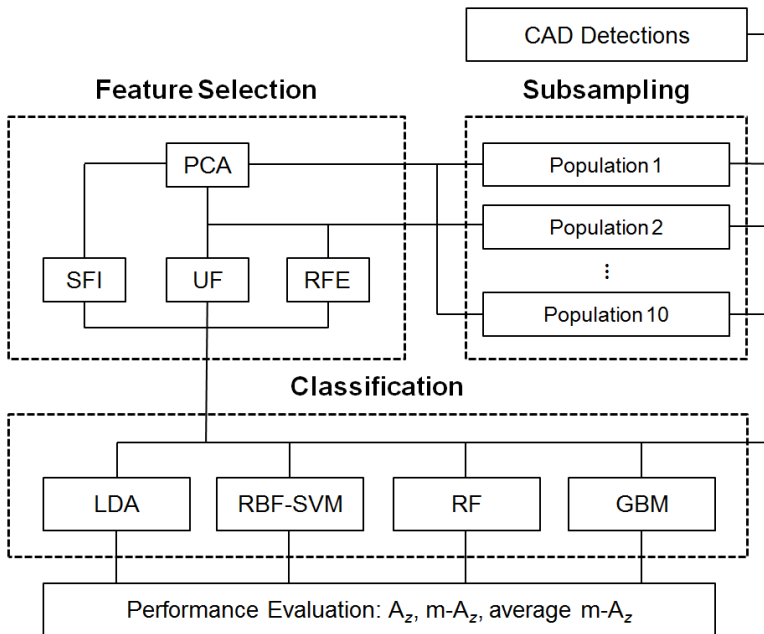


Fig. 1. Diagram illustrating the experiments of the study. The CAD detections were sampled randomly to construct 10 population samples. For each population, feature selection was performed using one of 5 methods (see Section 2.1) and classification was performed with or without feature selection using one of 4 methods (see Section 2.2). Performance evaluated was based on the area under the receiver operating characteristic curve for each population (A_z), average of A_z over the 10 populations (m- A_z), and average m- A_z over the feature selection methods or over the classifiers and each feature selection method (average m- A_z).

3 Results

The classification performance of the four classifiers is summarized in Tables 1 to 6. Tables 1 through 5 show the classifier performance with each of the five feature selection methods, whereas Table 6 shows the classifier performance without feature selection. Each row shows the result of an indicated population sample, whereas the columns indicate the average and standard deviation of the A_z value over three repeated 10-fold cross-validations of the indicated classifier. Bold numbers indicate the highest average of A_z for a sample. The bottom row shows the average and standard deviation of A_z over the 10 population samples.

If we consider the highest overall classifier performance in terms of the highest mean A_z ($m-A_z$) over the 10 subsampled populations, the ranking of classifiers varies according to the feature selection method. The performance was highest for the RBF-SVM without feature selection (0.800; Table 6), followed by the RF with RFE (0.799; Table 4), followed by RBF-SVM with the SFI of PCs (0.798; Table 1) and GBM with RFE (0.798; Table 4).

Also, the frequency at which a classifier outperformed the other classifiers in terms of the highest A_z for each of the 10 populations depended on the feature selection method. RBF-SVM outperformed the other classifiers most often with SFI of PCs (Table 1), with UF (Table 2), and without feature selection (Table 6). However, LDA outperformed the other classifiers with UF of PCs (70%; Table 3) and RFE of PCs (40%; Table 5). The RF and GBM classifiers outperformed the other classifiers with the RFE feature selection method (Table 4).

The robustness, or consistency, of a classifier, can be characterized by the average $m-A_z$ value that the classifier yields for the different feature selection methods. In this sense, RBF-SVM yielded the highest performance (0.795), followed by the RF (0.794), GBM (0.793), and LDA (0.786).

Table 1. Performance comparison of classifiers with feature selection by the SFI of PCs. Brackets indicate the standard deviation (SD) of A_z . The numbers in bold indicate the highest value of A_z among classifiers for each population sample.

Population	LDA	RBF-SVM	RF	GBM
1	0.780 [0.040]	0.796 [0.047]	0.789 [0.040]	0.778 [0.041]
2	0.789 [0.055]	0.796 [0.051]	0.794 [0.050]	0.780 [0.054]
3	0.810 [0.030]	0.821 [0.028]	0.814 [0.024]	0.812 [0.032]
4	0.789 [0.039]	0.808 [0.041]	0.802 [0.045]	0.790 [0.038]
5	0.794 [0.043]	0.790 [0.046]	0.786 [0.057]	0.789 [0.055]
6	0.806 [0.035]	0.800 [0.041]	0.793 [0.044]	0.798 [0.041]
7	0.789 [0.035]	0.804 [0.037]	0.790 [0.034]	0.799 [0.036]
8	0.786 [0.042]	0.790 [0.041]	0.785 [0.042]	0.783 [0.044]
9	0.776 [0.057]	0.789 [0.064]	0.792 [0.055]	0.781 [0.057]
10	0.772 [0.052]	0.790 [0.049]	0.791 [0.043]	0.789 [0.039]
Mean \pm SD	0.789 \pm 0.012	0.798 \pm 0.010	0.793 \pm 0.009	0.790 \pm 0.011

Table 2. Performance comparison of classifiers with feature selection by UF

Population	LDA	RBF-SVM	RF	GBM
1	0.784 [0.040]	0.782 [0.041]	0.790 [0.038]	0.781 [0.048]
2	0.785 [0.058]	0.793 [0.049]	0.796 [0.054]	0.798 [0.054]
3	0.809 [0.029]	0.816 [0.024]	0.815 [0.025]	0.807 [0.032]
4	0.778 [0.047]	0.799 [0.048]	0.793 [0.044]	0.797 [0.053]
5	0.779 [0.038]	0.785 [0.050]	0.778 [0.055]	0.783 [0.051]
6	0.790 [0.035]	0.791 [0.042]	0.776 [0.046]	0.771 [0.042]
7	0.789 [0.035]	0.811 [0.035]	0.792 [0.031]	0.797 [0.027]
8	0.772 [0.041]	0.791 [0.048]	0.794 [0.042]	0.790 [0.043]
9	0.764 [0.057]	0.791 [0.053]	0.802 [0.054]	0.796 [0.052]
10	0.756 [0.056]	0.800 [0.044]	0.780 [0.043]	0.784 [0.035]
Mean \pm SD	0.781 \pm 0.015	0.796 \pm 0.011	0.792 \pm 0.012	0.790 \pm 0.010

Table 3. Performance comparison of classifiers with feature selection by UF of PCs

Population	LDA	RBF-SVM	RF	GBM
1	0.796 [0.040]	0.796 [0.047]	0.792 [0.047]	0.785 [0.041]
2	0.805 [0.051]	0.799 [0.051]	0.800 [0.043]	0.791 [0.052]
3	0.819 [0.024]	0.812 [0.028]	0.816 [0.031]	0.815 [0.034]
4	0.806 [0.038]	0.798 [0.041]	0.791 [0.040]	0.803 [0.038]
5	0.797 [0.045]	0.769 [0.046]	0.768 [0.060]	0.779 [0.044]
6	0.804 [0.039]	0.784 [0.041]	0.783 [0.041]	0.790 [0.042]
7	0.795 [0.035]	0.786 [0.037]	0.784 [0.034]	0.795 [0.031]
8	0.789 [0.037]	0.782 [0.041]	0.792 [0.035]	0.788 [0.037]
9	0.782 [0.053]	0.780 [0.064]	0.775 [0.054]	0.783 [0.048]
10	0.765 [0.054]	0.779 [0.049]	0.772 [0.040]	0.776 [0.043]
Mean \pm SD	0.796 \pm 0.015	0.789 \pm 0.013	0.787 \pm 0.014	0.790 \pm 0.011

Table 4. Performance comparison of classifiers with feature selection by RFE

Population	LDA	RBF-SVM	RF	GBM
1	0.786 [0.041]	0.789 [0.041]	0.800 [0.040]	0.784 [0.044]
2	0.739 [0.054]	0.798 [0.047]	0.803 [0.050]	0.810 [0.052]
3	0.805 [0.030]	0.814 [0.025]	0.819 [0.024]	0.806 [0.028]
4	0.776 [0.047]	0.800 [0.048]	0.798 [0.045]	0.800 [0.046]
5	0.769 [0.048]	0.788 [0.050]	0.789 [0.057]	0.790 [0.053]
6	0.802 [0.038]	0.789 [0.042]	0.778 [0.044]	0.785 [0.040]
7	0.773 [0.034]	0.813 [0.038]	0.799 [0.034]	0.810 [0.027]
8	0.765 [0.042]	0.790 [0.047]	0.798 [0.042]	0.791 [0.041]
9	0.759 [0.063]	0.792 [0.053]	0.810 [0.055]	0.801 [0.051]
10	0.753 [0.054]	0.801 [0.044]	0.791 [0.043]	0.801 [0.039]
Mean \pm SD	0.773 \pm 0.021	0.797 \pm 0.010	0.799 \pm 0.011	0.798 \pm 0.010

Table 5. Performance comparison of classifiers with feature selection by RFE of PCs

Population	LDA	RBF-SVM	RF	GBM
1	0.800 [0.042]	0.799 [0.044]	0.799 [0.041]	0.787 [0.044]
2	0.800 [0.055]	0.780 [0.058]	0.802 [0.046]	0.773 [0.051]
3	0.817 [0.029]	0.809 [0.029]	0.815 [0.029]	0.814 [0.042]
4	0.796 [0.040]	0.796 [0.036]	0.805 [0.043]	0.804 [0.042]
5	0.807 [0.044]	0.796 [0.045]	0.788 [0.054]	0.785 [0.047]
6	0.811 [0.037]	0.794 [0.034]	0.796 [0.039]	0.817 [0.035]
7	0.810 [0.035]	0.792 [0.040]	0.773 [0.044]	0.798 [0.034]
8	0.797 [0.037]	0.759 [0.044]	0.804 [0.034]	0.797 [0.029]
9	0.772 [0.059]	0.782 [0.059]	0.789 [0.057]	0.792 [0.051]
10	0.752 [0.045]	0.793 [0.046]	0.801 [0.039]	0.804 [0.038]
Mean \pm SD	0.796 \pm 0.020	0.790 \pm 0.014	0.797 \pm 0.012	0.797 \pm 0.013

Table 6. Performance comparison of classifier without feature selection

Population	LDA	RBF-SVM	RF	GBM
1	0.785 [0.041]	0.792 [0.043]	0.797 [0.040]	0.783 [0.040]
2	0.785 [0.054]	0.800 [0.049]	0.801 [0.050]	0.794 [0.057]
3	0.804 [0.032]	0.815 [0.025]	0.815 [0.023]	0.805 [0.033]
4	0.782 [0.043]	0.809 [0.044]	0.798 [0.041]	0.799 [0.043]
5	0.784 [0.047]	0.792 [0.050]	0.784 [0.056]	0.788 [0.051]
6	0.798 [0.038]	0.799 [0.042]	0.776 [0.043]	0.780 [0.038]
7	0.784 [0.035]	0.813 [0.038]	0.799 [0.031]	0.809 [0.029]
8	0.776 [0.043]	0.791 [0.044]	0.794 [0.042]	0.790 [0.038]
9	0.772 [0.057]	0.794 [0.055]	0.803 [0.054]	0.799 [0.052]
10	0.763 [0.048]	0.798 [0.046]	0.791 [0.043]	0.798 [0.038]
Mean \pm SD	0.783 \pm 0.012	0.800 \pm 0.009	0.796 \pm 0.011	0.794 \pm 0.009

Similarly, we can also characterize the robustness of a feature selection method by calculation of the average $m-A_z$ value of the classifiers yielded by each selection method. In this sense, RFE of PCs yielded the highest performance (average of $m-A_z = 0.795$; Table 5), followed by the strategy without feature selection (0.793; Table 6), SFI of PCs (0.793; Table 1), RFE (0.792; Table 4), UF of PCs (0.791; Table 3), and UF (0.790; Table 2).

4 Discussion

The preliminary results of this pilot study indicate that the RBF-SVM classifier compares favorably with other state-of-the-art classifiers in the discrimination of TP and FP CAD detections in CTC. The feature selection method of RFE of PCs compares favorably with other feature selection methods. However, no single classifier could be considered optimal under all conditions, including the use of different population samples or different feature selection methods.

In this pilot study, we used balanced sets of TP and FP samples. In practice, CAD systems produce unbalanced samples with a large number of FP samples and relatively few TP samples. However, the use of balanced data sets for the purposes of constructing classifiers and for estimating classification accuracy would often be more convenient and faster. Further work is needed for establishing the effect of using balanced and unbalanced sets on the projected classification accuracy.

Computational demands can place constraints on the classification problem. Among feature selection methods, RFE can be considered as a rather slow method for calculation, whereas the calculation for PCA is quite fast. The UF method is faster than the RFE, but slower than the PCA method. Among classifiers, the construction of SVM and GBM classifiers is remarkably slower than that of RF and LDA classifiers. The relatively small differences of the performance results suggest that although the use of fast classifiers and feature selection methods may reduce classification accuracy over that of slower methods, the reduction in overall accuracy is not necessarily meaningful in a practical application.

Acknowledgements. This work was supported by grants from NIH/NCI CA095279, CA166816, and CA131718. We also thank Partners Research Computing for providing high-performance computing services.

References

1. Yoshida, H., Näppi, J.: Three-Dimensional Computer-Aided Diagnosis Scheme for Detection of Colonic Polyps. *IEEE Trans. Med. Imaging* 20, 1261–1274 (2001)
2. Yoshida, H., Masutani, Y., MacEneaney, P., Rubin, D.T., Dachman, A.H.: Computerized Detection of Colonic Polyps at CT Colonography on the Basis of Volumetric Features: Pilot Study. *Radiology* 222, 327–336 (2002)
3. Summers, R.M., Jerebko, A.K., Franaszek, M., Malley, J.D., Johnson, C.D.: Colonic Polyps: Complementary Role of Computer-Aided Detection in CT Colonography. *Radiology* 225, 391–399 (2002)
4. Näppi, J., Yoshida, H.: Automated Detection of Polyps with CT Colonography: Evaluation of Volumetric Features for Reduction of False-Positive Findings. *Acad. Radiol.* 9, 386–397 (2002)
5. Näppi, J., Yoshida, H.: Feature-Guided Analysis for Reduction of False Positives in CAD of Polyps for Computed Tomographic Colonography. *Med. Phys.* 30, 1592–1601 (2003)
6. Suzuki, K., Zhang, J., Xu, J.: Massive-Training Artificial Neural Network Coupled with Laplacian-Eigenfunction-Based Dimensionality Reduction for Computer-Aided Detection of Polyps in CT Colonography. *IEEE Trans. Med. Imaging* 29, 1907–1917 (2010)
7. Näppi, J., Nagata, K.: Sources of False Positives in Computer-Assisted CT Colonography. *Abdom. Imaging* 36, 153–164 (2011)
8. Yoshida, H., Näppi, J.: CAD in CT Colonography without and with Oral Contrast Agents: Progress and Challenges. *Comput. Med. Imaging Graph.* 31, 267–284 (2007)
9. Näppi, J.J., Regge, D., Yoshida, H.: Ensemble Detection of Colorectal Lesions for CT Colonography. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) *Abdominal Imaging 2011*. LNCS, vol. 7029, pp. 60–67. Springer, Heidelberg (2012)
10. Nadler, M., Smith, E.P.: *Pattern Recognition Engineering*. Wiley, New York (1993)

11. Martinez, A.M., Kak, A.C.: PCA versus LDA. *IEEE Trans. Pattern Anal. Mach. Intell.* 23, 228–233 (2001)
12. Mankiewicz, R.: *The Story of Mathematics*, p. 158. Princeton University Press (2001)
13. Guyon, I., Weston, J., Barnhill, S., Vapnik, V.: Gene Selection for Cancer Classification Using Support Vector Machines. *Mach. Learn.* 46, 389–422 (2002)
14. Schölkopf, B., Burges, C.J.C., Smola, A.J. (eds.): *Advances in Kernel Methods: Support Vector Learning*. MIT Press, Cambridge (1999)
15. Breiman, L.: Random Forests. *Mach. Learn.* 45, 5–32 (1999)
16. Svetnik, V., Liaw, A., Tong, C., Wang, T.: Application of Breiman's Random Forest to Modeling Structure-Activity Relationships of Pharmaceutical Molecules. In: Roli, F., Kittler, J., Windeatt, T. (eds.) *MCS 2004. LNCS*, vol. 3077, pp. 334–343. Springer, Heidelberg (2004)
17. Hastie, T., Tibshirani, R., Friedman, J.H.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, Heidelberg (2001)
18. Friedman, J.H.: Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Statist.* 29, 1189–1232 (2001)
19. Näppi, J., Frimmel, H., Dachman, A., Yoshida, H.: A New High-Performance CAD Scheme for the Detection of Polyps in CT Colonography. In: *SPIE Medical Imaging 2004*, vol. 5370, pp. 839–848 (2004)
20. Näppi, J., Yoshida, H., Regge, D.: Accuracy of Computer-Aided Detection in Detecting Flat Lesions in CT Colonography. *Insights into Imaging* 3(S1), 334 (2012)

Piecewise Structural Diffusion Defined on Shape Index for Noise Reduction in Dual-Energy CT Images

Wenli Cai¹, June-Goo Lee^{1,2}, Da Zhang¹, Christina Piel^{1,3}, and Hiroyuki Yoshida¹

¹ 3D Imaging Research, Department of Radiology,
Massachusetts General Hospital and Harvard Medical School,
25 New Chardon St., Suite 400C, Boston, Massachusetts 02114, USA
{cai.wenli,yoshida.hiro}@mgh.harvard.edu

² Department of Radiology, University of Pittsburgh, 3362 Fifth Ave.
FARP/Imaging Research, Pittsburgh, Philadelphia 15213, USA

³ DInstitut für Medizinische Physik und Strahlenschutz, University of
Applied Sciences in Giessen, Wiesenstraße 14, 35390 Giessen, Germany

Abstract. The increasing radiation dose in dual-energy CT (DE-CT) scanning due to the double exposures at 80 kVp and 140 kVp is a major concern in the application of DE-CT. This paper presents a novel image-space denoising method, called *piecewise structural diffusion* (PSD), for the reduction of noise in low-dose DE-CT images. Three principle structures (plate, ridge, and cap) and their corresponding diffusion tensors are formulated based on the eigenvalues of a Hessian matrix. The local diffusion tensor that is piecewise-defined on the domain of shape index is composed by a linear combination of two diffusion tensors of the associated principle structures. A single diffusion tensor calculated from the fused DE-CT image is applied to both high- and low-energy images. In the DE-CT colon phantom study, we demonstrated that DE-CT images filtered by PSD yielded the similar image quality with half of radiation doses.

Keywords: Noise reduction, dual-energy CT, dual-energy CT colonography.

1 Introduction

The *dual-energy CT* (DE-CT) theory was first introduced by Alvarez and Macovski more than three decades ago [1]. With the recent technical advances, DE-CT became widely available in clinical practice. It provides an effective means for estimation of material composition by an analysis of two X-ray attenuation values acquired simultaneously at two photon energies (such as 80 kVp and 140 kVp). This ability of material decomposition has generated various new clinical applications that are otherwise unavailable in conventional single-energy CT [2].

Although DE-CT is highly promising, the potential of increased radiation dose, in the vicinity of one and half to two times of the dose of a routine single-energy CT examination due to the double exposures at 80 kVp and 140 kVp, is a major concern in the clinical application of DE-CT, because the potential risk of radiation-induced cancer due to the increasing radiation dose may not be negligible. Therefore, it is

imperative to lower the radiation dose of DE-CT to the level that is comparable to conventional single-energy CT without sacrificing the image quality.

Reduction of radiation dose (such as by lowering the tube current) may cause high image noise, low image contrast, and increased artifacts [3]. Major CT manufactures have been developing solutions to reducing radiation dose for DE-CT scanning, such as the automatic dose exposure control [4] and iterative reconstruction methods [5]. These methods are based on either the scanner hardware or the projection data (sinogram). An alternative solution is to apply noise reduction filters to reconstructed DE-CT images acquired with a low radiation dose, i.e., image-space denoising filters.

The purpose of this study is to develop a novel image-space DE-CT noise reduction filter called *piecewise structural diffusion* (PSD), and to evaluate the performance and the potential to reducing radiation dose, by using nine DE-CT datasets of a colon phantom scanned at various radiation doses.

The main contributions of the study are (1) PSD can deal with multiple structures in one diffusion process, and (2) a diffusion function is piecewise-defined on shape index [6-9]. Conventional diffusion filters works on only one type of structures, such as either tube or surface. We defined multiple principle structures (plate, ridge, and cap) based on anatomic features of interest, and used the scalar value of shape index to characterize a local 3D topologic shape. A local structure and its diffusion function were piecewise-defined on shape index. Our secondary contribution is to apply a single diffusion tensor to both high- and low-energy images in DE-CT, which takes advantages of both images for estimating a reliable local structure. We developed and demonstrated PSD filter in DE-CT colonography images.

2 Methods

2.1 Previous Work

Noise reduction filters have been developed and assessed for lowering image noise and improving image quality on CT images for decades [10]. Perona and Malik introduced the anisotropic diffusion process [11], in which smoothing is formulated mathematically as a diffusive process,

$$\frac{\partial}{\partial x} u(x, t) = \text{div}(D(x, t)\nabla u(x, t)), \quad (1)$$

where div is the divergence operator, ∇u is the gradient of the image and the diffusion strength is controlled by the diffusion tensor $D(x, t)$. $u(x, t)$ is the image intensity at spatial coordinate x and iteration step t .

In general, the diffuse tensor $D: R^{3 \times 3} \rightarrow R^{3 \times 3}$ is a matrix that enforces the directional preference of the diffusion along the principal directions. The diffusion tensor must satisfy C^∞ continuous, symmetric, and positive definite. Therefore, D is a symmetric positive definite matrix for a nonlinear anisotropic diffusion.

Conventional diffusion filters deals with one type of structures, such as edge-enhancing diffusion that enhances surface (edge) and coherent-enhancing diffusion

that enhances tube (line) [12]. However, anatomic structures may have different shapes in addition to surface and tube. In the context of CT colonography (CTC), colon contains surface shape for colonic walls, ridge shape for colonic folds, and cap shape for colonic polyps. Therefore, it is essential that a diffusion process can deal with multiple shapes in order to preserve different anatomic structures interested in images.

2.2 Principle Structure

Let $u(x)$ denote the CT value at a point $x \in R^3$ in a CT volume. The local morphological structure of $u(x)$ can be represented by a Hessian matrix $\nabla_{\sigma}^2 u(x)$, which is calculated by convolving $u(x)$ with the second Gaussian derivatives of a standard deviation σ . Let $|\lambda_1(x)| \leq |\lambda_2(x)| \leq |\lambda_3(x)|$ be the eigenvalues of $\nabla_{\sigma}^2 u(x)$ with corresponding eigenvectors of $e_1(x), e_2(x), e_3(x)$. A local morphologic structure can be characterized with a combination of the eigenvalues, called *eigenvalue signatures* [13].

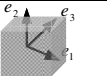
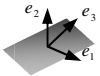
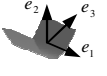

In general, a 3D diffuse tensor D can be defined as,

$$D(\nabla_{\sigma}^2 u(x)) = V(x)\Lambda(x)V(x)^T, \tag{2}$$

where $V(x) = [e_1(x), e_2(x), e_3(x)]$, $\Lambda(x) = \text{diag}[\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3]$, and $\tilde{\lambda}_i (0 < \tilde{\lambda}_i \leq 1)$ is a positive-definite function of λ_i .

Table 1 illustrates the eigenvalue signatures and diffusion functions of different types of structures. Morphologically, if all eigenvalues are close to zero, there is no structure, i.e., it is a homogeneous region. A plate-like structure (such as colonic walls) has eigenvalue signature of $\lambda_1, \lambda_2 \approx 0; \lambda_3 \gg 0$. A ridge-like structure (such as colonic folds) has eigenvalue signature of $\lambda_1 \approx 0; \lambda_2, \lambda_3 \gg 0$. A cap-like structure (such as colonic polyps) has eigenvalue signature of $\lambda_1, \lambda_2, \lambda_3 \gg 0$. The sign of λ_i corresponds to convex/bright shapes ($\lambda_i \leq 0$) or concave/dark shape ($\lambda_i \geq 0$).

Table 1. Eigenvalue signatures and diffusion functions of principle structures

Morphological Category	Eigenvalue Signature	Diffusion Function
No structure (homogeneous) 	$\lambda_1, \lambda_2, \lambda_3 \approx 0$	(1,1,1)
Plate (colonic wall) 	$\lambda_1, \lambda_2 \approx 0;$ $\lambda_3 \gg 0$	$(1,1, \tilde{\lambda}_3^{plate})$
Ridge (colonic fold) 	$\lambda_1 \approx 0;$ $\lambda_2, \lambda_3 \gg 0$	$(1, \tilde{\lambda}_2^{Ridge}, \tilde{\lambda}_3^{Ridge})$
Cap (colonic polyp) 	$\lambda_1, \lambda_2, \lambda_3 \gg 0$	$(\tilde{\lambda}_1^{Cap}, \tilde{\lambda}_2^{Cap}, \tilde{\lambda}_3^{Cap})$

Because diffusion tensor D in equation (2) is required to be positive definite, all eigenvalues of D are positives. This also indicates that the diagonal element of $\Lambda(x)$ is positive definite, i.e., $\tilde{\lambda}_i^{ST} > 0$. Therefore, we assume, without loss of generality, that $\tilde{\lambda}_i^{ST}$ is a function of $|\lambda_i|$.

2.3 Diffusion Functions of Principle Structures

A diffusion function $\tilde{\lambda}_i^{ST}$ of a principle structure is defined as a function of eigenvalues. We define $\tilde{\lambda}_3^{ST}$ as a function of λ_3 :

$$\tilde{\lambda}_3^{ST}(\lambda_3) = \begin{cases} 1 & ; |\lambda_3| \leq H \\ 1 - \exp\left(\frac{-C}{(|\lambda_3| - H)^2}\right) & ; \text{otherwise,} \end{cases} \quad (3)$$

where H is a constant of eigenvalue to indicate whether there is a structure at point x (in the case of $|\lambda_3(x)| > H$), and $C = 2.0$ is a threshold parameter.

The diffusion functions $\tilde{\lambda}_1^{ST}$ and $\tilde{\lambda}_2^{ST}$ for a principle structure are calculated based on the ratios of $\frac{\lambda_1}{\lambda_3}$ and $\frac{\lambda_2}{\lambda_3}$:

$$\tilde{\lambda}_i^{ST}(\lambda_i) = \tilde{\lambda}_3^{ST} + (1 - \tilde{\lambda}_3^{ST}) \cdot \exp\left(-k \cdot \left(\frac{\lambda_i}{\lambda_3}\right)^2\right), \quad (4)$$

where $i = 1, 2$, and $k = 10$ is a threshold parameter.

In order to stop the diffusion at edges near zero-crossing, a weight function h is defined by use of the directional second derivative u_{nn} along the normal \mathbf{n} :

$$h(u_{nn}) = 1 - 0.9 \left(\frac{u_{nn}}{U}\right)^2, \quad (5)$$

where $U = 5$ is a threshold parameter. Note that $h = 0.0$ at zero-crossings.

The final diffusion function is formulated by $h(u_{nn}) * \tilde{\lambda}_i^{ST}(\lambda_i)$.

2.4 Shape Index

A shape index $SI(x)$ characterizes the local 3D topologic shape at a point $x \in R^3$ by use of the two principal curvatures κ_{max} and κ_{min} [6],[11],[12],[13]. Because the sign of a curvature defines the bending direction, i.e., inner or outer, a positive SI represents a convex shape, whereas a negative SI represents a concave shape. Under the condition of $\lambda_i \geq 0$, the shape index is defined by absolute value of curvatures as follows:

$$SI(x) = \frac{2}{\pi} \arctan\left(\frac{|\kappa_{max}| + |\kappa_{min}|}{|\kappa_{max}| - |\kappa_{min}| + \delta}\right), \quad (6)$$

where $\delta = 0.001$.

Fig. 1 illustrates that, by using equation (6), it is possible to determine to which of the three topologic shapes, plate, ridge, or cap, a point belongs based on a single value

of SI . The shape index is normalized between 0 and 1. Thus, point x that belongs to the plate shape has SI values around 0; ridge, around 0.5; and cap, around 1.0. The transition from one topologic shape to another occurs continuously.

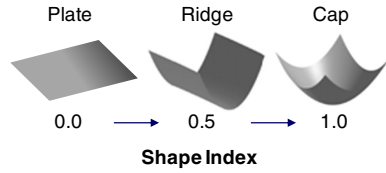


Fig. 1. Illustration of the relationship between the values of the shape index (equation (6)) and the corresponding shapes

2.5 Piecewise Structural Diffusion

Under the condition of $|\lambda_1| + |\lambda_2| + |\lambda_3| > 0$, the local shape at a point x can be represented as a combination of two principle structures, called *piecewise-defined structure*. Point x that has a combination of plate and ridge has $SI(x) < 0.5$; otherwise, it has a combination of ridge and cap. Therefore, a diffusion function is a piecewise function defined on $SI(x)$, called *piecewise structural diffusion* (PSD), as shown in equation (7).

$$\tilde{\lambda}_i = \begin{cases} (1 - w) \cdot \tilde{\lambda}_i^{Plate} + w \cdot \tilde{\lambda}_i^{Ridge}; & (0.0 \leq SI < 0.5) \\ (1 - w) \cdot \tilde{\lambda}_i^{Ridge} + w \cdot \tilde{\lambda}_i^{Cap}; & (0.5 \leq SI < 1.0), \end{cases} \quad (7)$$

where w is the weighting factor, which is in the range of $[0.0, 1.0]$.

Suppose that $P1$ and $P2$ are two principle structures at point x so that $SI^{P2} > SI^{P1}$. A simple form of w is a linear function of SI , as shown in equation (8).

$$w(x) = \frac{SI(x) - SI^{P1}}{SI^{P2} - SI^{P1}}. \quad (8)$$

2.6 Dual-Energy CT Diffusion

In DE-CT scanning, there are two CT values at a point x , $u_L(x)$ and $u_H(x)$, which are the CT values at low (80 kVp) and high (140 kVp) energies, respectively. In general, 80 kVp images provide greater contrast than that of 140 kVp, but its image quality is limited due to the increased noise. Because u_L and u_H are CT values at different photon energies of the same anatomic structures, the underlying diffusion structures are essentially the same for both images. Therefore, we applied a single diffusion tensor to both images as illustrated in Fig. 2.

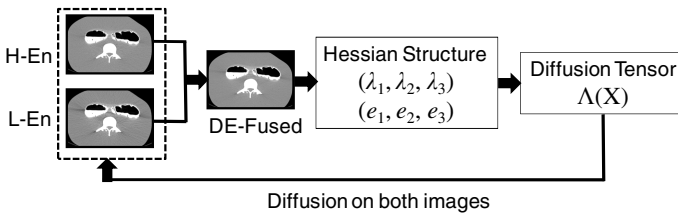


Fig. 2. Illustration of dual-energy CT diffusion iteration

The single diffusion tensor, $\Lambda(x)$, for both u_L and u_H was calculated by use of the fused images. Conventionally, high-energy and low-energy images are fused by a linear blending function: a mixing of 30% 80 kVp and 70% 140 kVp ($0.3 \cdot u_L(x) + 0.7 \cdot u_H(x)$).

3 Experiments and Results

We used a custom-ordered anthropomorphic colon phantom (Phantom Laboratory, Salem, NY) that was made of a urethane-mix material with a CT value of -100 ± 10 HU (120 kVp), and cast around simulated bony structures of a pelvis, 2 femurs, and 3 lumbar vertebrae. We installed a question-mark-shaped simulated colon with a length of approximately 50 cm and diameters of 3.5 ~ 4.0 cm. To simulate the semi-solid fecal materials in fecal-tagging CTC, we filled the colon phantom, prior to imaging, with 300 ml of simulated non-cathartic tagged fecal residues that were a mixture of aqueous fiber (psyllium), ground foodstuff (cereal), and non-ionic iodinated contrast agent (Omnipaque iohexol, GE Healthcare) at a concentration of 40 mg/ml.

The phantom was scanned on a DE-CT scanner (SOMATON Definition Flash, Siemens Healthcare) with two different photon voltages: 80 kVp and 140 kVp. We applied nine tube current settings, 12 ~ 74 mAs for 140 kVp and 60 ~ 370 mAs for 80 kVp, to test image noise at different radiation doses. The radiation dose index (CTDIvol) ranged from 2.79 to 16.77 mGy. For all scanning, we applied the soft tissue reconstruction algorithm and a 0.625 mm slice reconstruction interval. In total, we generated nine DE-CT datasets (512x512x537) for testing.

We used the same parameter settings for all of our denoising experiments. With a unit grid spacing in all dimensions, we empirically set $\Delta t = 0.125$ for a stable numeric iteration of the diffusion equation (1). The number of iterations in the diffusion process was set to 50. The standard deviation of Gaussian derivatives for calculation of Hessian matrix and eigenvalues was set to $\sigma = 1.0$.

For the assessment of the image quality after noise reduction, two groups of image quality metrics were evaluated:

- To measure the noise level of the images: the mean (μ) and standard deviation (SD) of the images.
- To measure the local structure similarity: quality index of local variance (QILV) [11] (value range [0.0, 1.0]) referred to the images scanned at the highest radiation dose.

Table 2 lists the measurements of mean SD (MSD) and QILV for both original and denoised images at 140 kVp and 80 kVp. By comparison of the shaded cells in the table, we observed that the PSD filter can reduce the radiation dose approximately 50% for the 140 kVp images and more than 50% for the 80 kVp images.

Fig. 3 demonstrates an example of the effect of PSD on the reduction of radiation dose. The images at the upper row are the original images scanned at a CTDIvol of 13.65 mGy, and the lower row shows the PSD-denoised images that were scanned at a CTDIvol of half of the dose, i.e., 6.83 mGy. Both groups of images have the same

Table 2. Comparison of mean standard deviation (MSD) and quality index of local variance (QILV) in both 140 kVp and 80 kVp images

Dose (mGy)	140 kVp Images				80 kVp Images			
	Original		Denoised		Original		Denoised	
	MSD	QILV	MSD	QILV	MSD	QILV	MSD	QILV
2.79	69.2	0.86	49.0	0.97	97.1	0.93	63.0	0.97
3.50	66.5	0.90	45.9	0.98	85.1	0.95	52.8	0.98
4.65	55.9	0.94	36.2	0.99	79.0	0.97	48.4	0.98
6.83	46.9	0.98	26.6	0.99	69.2	0.98	39.0	0.99
9.10	41.0	0.99	21.0	0.99	67.4	0.99	36.2	0.99
11.34	38.0	0.99	19.5	0.98	61.4	0.99	32.1	0.98
13.65	35.6	0.99	17.8	0.98	51.2	0.99	25.9	0.98
15.92	30.8	0.99	14.3	0.98	48.4	0.99	26.3	0.98
16.77	30.4	1.00	13.6	0.99	48.0	1.00	25.7	0.99

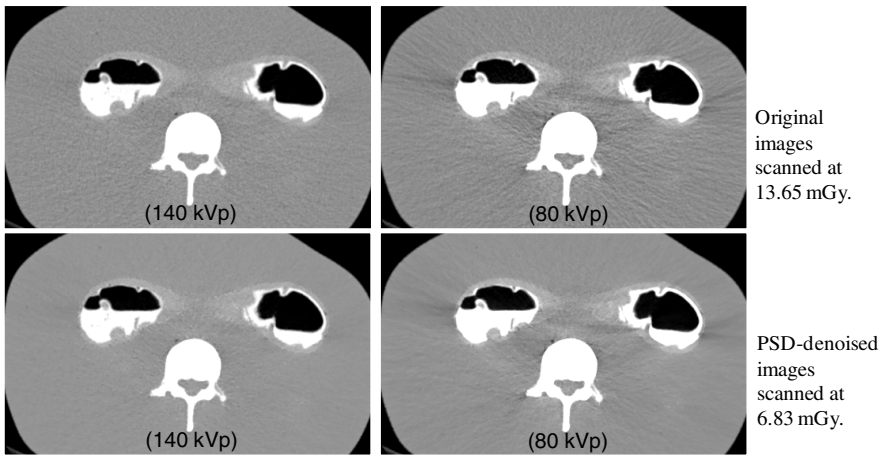


Fig. 3. An example of the effect of PSD filter on the reduction of radiation dose. The upper row shows the original images scanned at 13.65 mGy with MSD of 35.6 (140 kVp) and 51.2 (80 kVp). The lower row shows the denoised images scanned at 6.83 mGy with MSD of 26.6 (140 kVp) and 39.0 (80 kVp). Both groups have the same QILV of 0.99.

QILV values, i.e., the same level of structure similarity referred to the images scanned at the highest dose. However, the MSD of the images scanned at 13.65 mGy (upper row) were 35.6 (140 kVp) and 51.2 (80 kVp), whereas the MSD of the PSD-denoised images scanned at 6.83 mGy (lower row) were 26.6 (140 kVp) and 39.0 (80 kVp), respectively. We observed that the half-dose DE-CT images have less noise after the application of PSD filter than that of the full-dose images.

We compared the performance of PSD with six other denoising methods: binomial filter (BF), Gaussian filter (GF), curvature anisotropic diffusion (CAD), curvature flow (CF), gradient anisotropic diffusion (GAD), and min/max curvature diffusion (MMCD) by use of the recommended parameters [14]. We compared the SDs before

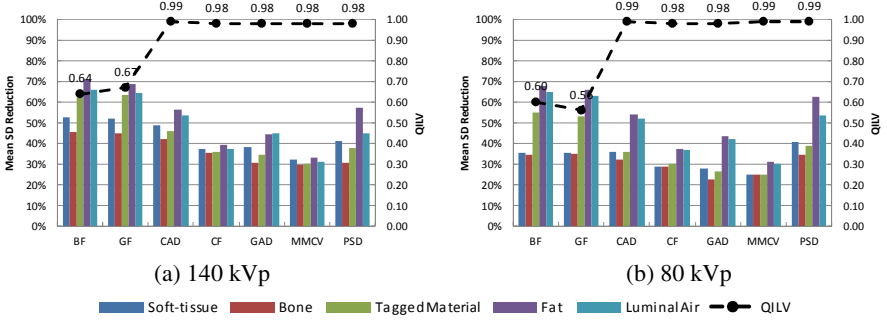


Fig. 4. Comparisons of seven denoising algorithms: binomial filter (BF), Gaussian filter (GF), curvature anisotropic diffusion (CAD), curvature flow (CF), gradient anisotropic diffusion (GAD), min/max curvature diffusion (MMCD), and the proposed PSD, in (a) 140 kVp images and (b) 80 kVp images. The left vertical axis is the mean SD reduction: $R_{SD} = (SD_{source} - SD_{filtered})/SD_{source} \cdot 100\%$, which measures the performance on noise reduction. The color bars show the mean R_{SD} in five materials. The right vertical axis is the mean QILV. The dashed line displays the mean values of QILV of the filtered images referred to the highest dose images.

and after application of seven denoising algorithms on five materials: soft-tissue, bone, tagged materials, fat, and luminal air. The noise reduction rate is define as $R_{SD} = (SD_{source} - SD_{filtered})/SD_{source} \cdot 100\%$. Mean R_{SD} values of nine datasets are compared by the color bars in Fig. 4 for both 140 kVp and 80 kVp images. In addition, we measured the QILV referred to the highest dose images to compare the effect of preservation of structural details after denoising. We observed that, although BF and GF had a high noise reduction rate, their QILV values were low compared to the other five diffusion filters. These results show that both BF and GF removed image noise by losing image details, which tend to blur the images. On the other hand, five diffusion filters showed better performance on detail preservation as indicated by the high QILV values. Among five diffusion filters, PSD performed better on noise reduction than other diffusion filters, especially in the low-energy images.

4 Conclusions

Image-space noise reduction filter is an alternate means for lowering the radiation dose in DE-CT. This paper presents a novel anisotropic diffusion filter called PSD that used a piecewise structural diffuse tensor for preserving multiple types of structures in DE-CT images. The proposed PSD filter was evaluated by use of nine DE-CT datasets of a colon phantom scanned at various radiation doses. The results demonstrated that the application of PSD may reduce more than half of the radiation dose in DE-CT scanning. In addition, PSD outperformed other diffusion filters on noise reduction in DE-CT images, especially in low-energy images. Further clinical studies will be conducted to demonstrate the performance of PSD filter on low-dose DE-CTC-based colon cancer screening and other clinical applications such as the electronic cleansing [15-17].

Acknowledgements. The project described was partly supported by Research Scholar Grant RSG-11-076-01-CCE from the American Cancer Society, and by Grant Numbers R01CA095279, R01CA131718, and R03CA156664 from National Cancer Institute (NCI).

References

1. Alvarez, R.E., Macovski, A.: Energy-Selective Reconstructions in X-Ray Computerised Tomography. *Phys. Med. Biol.* 21, 733–744 (1976)
2. Johnson, T., Fink, C., Schönberg, S.O., Reiser, M.F.: *Dual Energy CT in Clinical Practice* (2011)
3. Kalra, M.K., Maher, M.M., Toth, T.L., Hamberg, L.M., Blake, M.A., Shepard, J.A., Saini, S.: Strategies for CT Radiation Dose Optimization. *Radiology* 230, 619–628 (2004)
4. Murazaki, H., Funama, Y., Sugaya, Y., Miyazaki, O., Tomiguchi, S., Awai, K.: Optimal Setting of Automatic Exposure Control Based on Image Noise and Contrast on Iodine-Enhanced CT. *Acad. Radiol.* (2012)
5. Silva, A.C., Lawder, H.J., Hara, A., Kujak, J., Pavlicek, W.: Innovations in CT Dose Reduction Strategy: Application of the Adaptive Statistical Iterative Reconstruction Algorithm. *AJR Am. J. Roentgenol.* 194, 191–199 (2010)
6. Koenderink, J.J.: *Solid shape*. MIT Press, Cambridge (1990)
7. Yoshida, H., Näppi, J.: Three-Dimensional Computer-Aided Diagnosis Scheme for Detection of Colonic Polyps. *IEEE Trans. Med. Imaging* 20, 1261–1274 (2001)
8. Yoshida, H., Näppi, J., MacEneaney, P., Rubin, D.T., Dachman, A.H.: Computer-Aided Diagnosis Scheme for Detection of Polyps at CT Colonography. *Radiographics* 22, 963–979 (2002)
9. Yoshida, H., Masutani, Y., MacEneaney, P., Rubin, D.T., Dachman, A.H.: Computerized Detection of Colonic Polyps at CT Colonography on the Basis of Volumetric Features: Pilot Study. *Radiology* 222, 327–336 (2002)
10. Alvarez, R.E., Stonestrom, J.P.: Optimal Processing of Computed Tomography Images using Experimentally Measured Noise Properties. *J. Comput. Assist. Tomogr.* 3, 77–84 (1979)
11. Perona, P., Malik, J.: Scale-Space and Edge Detection using Anisotropic Diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 12, 629–639 (1990)
12. Weickert, J., Heidelberglaan, E.: *A Review of Nonlinear Diffusion Filtering* (1997)
13. Sato, Y., Westin, C., Bhalerao, A., Nakajima, S., Shiraga, N., Tamura, S., Kikinis, R.: Tissue Classification Based on 3D Local Intensity Structures for Volume Rendering. *IEEE Trans. Visualization and Computer Graphics* 6, 160–180 (2000)
14. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: *The ITK Software Guide*, 2nd edn. Kitware Inc. (2005)
15. Cai, W., Lee, J.G., Zalis, M.E., Yoshida, H.: Mosaic Decomposition: An Electronic Cleansing Method for Inhomogeneously Tagged Regions in Noncathartic CT Colonography. *IEEE Trans. Med. Imaging* 30, 559–574 (2011)
16. Cai, W., Yoshida, H., Zalis, M.E., Näppi, J.J., Harris, G.J.: Informatics in Radiology: Electronic Cleansing for Noncathartic CT Colonography: A Structure-Analysis Scheme. *Radiographics* 30, 585–602 (2010)
17. Cai, W., Zalis, M.E., Näppi, J., Harris, G.J., Yoshida, H.: Structure-Analysis Method for Electronic Cleansing in Cathartic and Noncathartic CT Colonography. *Med. Phys.* 35, 3259–3277 (2008)

A Supervised Learning Based Approach to Detect Crohn's Disease in Abdominal MR Volumes

Dwarikanath Mahapatra¹, Peter Schueffler¹, Jeroen A.W. Tielbeek², Joachim M. Buhmann¹, and Franciscus M. Vos^{2,3}

¹ Department of Computer Science, ETH Zurich, Switzerland
dwarikanath.mahapatra@inf.ethz.ch

² Department of Radiology, Academic Medical Center, The Netherlands

³ Quantitative Imaging Group, Delft University of Technology, The Netherlands

Abstract. Accurate diagnosis of Crohn's disease (CD) has emerged as an important medical challenge. Because current Magnetic resonance imaging (MRI) analysis approaches rely on extensive manual segmentation for an accurate analysis, we propose a method for the automatic identification and localization of regions in abdominal MR volumes that have been affected by CD. Our proposed approach will serve to augment results from colonoscopy, the current reference standard for CD diagnosis. Intensity statistics, texture anisotropy and shape asymmetry of the 3D regions are used as features to distinguish between diseased and normal regions. Particular emphasis is laid on a novel entropy based asymmetry calculation method. Experiments on real patient data show that our features achieve a high level of accuracy and perform better than two competing methods.

Keywords: Crohn's disease, supervised learning, magnetic resonance imaging.

1 Introduction

Inflammatory bowel diseases (IBDs) constitute one of the largest healthcare problems in the Western world afflicting over 1 million European citizens. Out of these, nearly 700,000 suffer from Crohn's disease (CD). Crohn's disease is a autoimmune IBD that may affect any part of the gastrointestinal tract causing abdominal pain, diarrhea, vomiting or weight loss. Assessment of CD severity is essential to determine the therapeutic strategy. Currently, the reference standard for diagnosis relies on results of colonoscopy and biopsy samples [13]. However, the procedure is invasive, requires extensive bowel preparation, and gives information only on superficial abnormalities. Therefore it is beneficial to have a non-invasive approach to detect CD.

Scope of Our Work: This paper proposes a method to *detect* and *localize* CD afflicted regions from input abdominal magnetic resonance (MR) volumes without an explicit segmentation of the bowel wall. Our method will serve as a tool

to assist clinicians, reduce reliance on colonoscopy and help in rapid diagnosis of CD. We extract features from 3D volumes of interest (VOI) and use them to detect diseased regions. Intensity statistics, texture anisotropy and curvature asymmetry were used as features to discriminate between diseased, normal and background (normal non-intestine) regions. Also, higher order intensity statistics like kurtosis and skewness that cannot be processed by the human visual system (HVS) [9] are used for disease classification. Anisotropy of features is a measure of a region’s asymmetry and has been used for identifying tumorous regions [11]. We propose a novel entropy based method to calculate the texture anisotropy and curvature asymmetry of a VOI. Experimental results show a high degree of accuracy in detecting CD. This paper makes the following contributions: 1) an entropy based approach to calculate anisotropy and asymmetry of a region is proposed; 2) it is combined with higher order image statistics to identify CD affected areas in abdominal MRI. We describe our method in Section 2, present results in Section 3 and conclude with Section 4.

Related Work on Disease Classification: Several drawbacks of colonoscopy like invasiveness, procedure related discomfort and risk of bowel perforation has led to the exploration of imaging techniques like sonography, computed tomography (CT) and MRI to assess extension and severity of IBDs [17]. MRI has the potential to overcome limitations of sonography (gas interposition) and CT (exposure to ionising radiations) because of high tissue contrast, lack of ionising radiations and lower incidence of adverse events related to intravenous contrast. Rimola et al. in [17] determined that rate of contrast enhancement and bowel wall thickness relate to the severity of endoscopically active CD. However, its reliance on explicit segmentation of the bowel wall and extensive manual scoring limits its effectiveness.

There does not exist abundant research on image analysis of abdominal MRI to identify Crohn’s disease, although [4] use dynamic contrast enhanced (DCE) MRI for identifying colorectal cancer and [18] deal with ulcerative colitis. Previous research has addressed the tasks of identifying abnormal regions in different applications like glaucoma diagnosis [7], whole body MR scans [15], cardiac coronary stenoses [10] and Alzheimer’s disease [21] to name a few.

2 Methods

We employ a two stage classification approach where a test sample is first classified as either background or intestine. If the initial classification is intestine, then the sample is further classified as either diseased or normal. Manual annotations by experts indicating diseased, normal and background regions in 3D MR volumes were available. For every such labeled voxel we extract features (intensity statistics, texture anisotropy and shape asymmetry) from its $35 \times 35 \times 5$ neighborhood for training the classifiers.

Intensity and Texture Features: Since a simple visual examination of T1 MRI does not provide sufficient information to identify diseased areas, radiologists rely

on the results from different tests (like colonoscopy and biopsy), and imaging protocols like MR-T1, MR-T2 and DCE-MRI. It is common in MR images to have regions that do not form distinct spatial patterns but differ in their third order statistics, e.g. boundaries of some malignant tumours are diffuse and invisible to the naked eye [16]. We propose to investigate features that are not discernible by the human eye but may provide discriminating features for our task. In addition to mean and variance we calculate third and fourth order intensity statistics (like skewness and kurtosis) in a voxel’s neighborhood.

In [11] texture anisotropy was used as a feature to identify tumorous regions in brain images. Normal tissues exhibit a regular pattern in their appearance while diseased regions show areas of asymmetry where the tissues have been affected by the progression of disease. We aim to exploit this characteristic to discriminate between diseased and normal regions in a VOI. Texture maps of the VOI were obtained using orthonormal Gabor filters in the $x - y$ and $y - z$ directions as described in [20] because 3D texture filters are computationally very expensive. Gabor filters conform to the receptive field properties of cortical cells, capture rich visual properties like spatial frequency characteristics and orientation, and are robust to noise by incorporating Gaussian smoothing. Texture maps are obtained along six directions ($0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$) for each slice.

While anisotropy in [11] was calculated using local gradient differences and gray level dependence histograms, we use entropy to measure anisotropy. Texture maps of slices in the $x - y$ plane are divided into 9 equal parts corresponding to 9 sectors of a circle, and entropy determined for each sector. A higher entropy value indicates wider distribution of texture values (hence high anisotropy), while low entropy indicates lower anisotropy. The texture anisotropy for sector r is

$$Tex_{Anisotropy}^r = - \sum_{tex} p_{tex}^r \log p_{tex}^r. \quad (1)$$

p_{tex}^r denotes the probability of distribution of texture values in sector r . We use a different approach to calculate anisotropy of images in the $y - z$ plane. As there are 5 slices in the z dimension, there are not enough samples (only $35 \times 5 = 175$) to calculate a reliable measure of entropy. Instead we calculate the entropy for the whole slice and use it in the feature vector. Thus the number of texture features are 80 (9 entropy values from each of 5 slices in the $x - y$ plane, and 1 entropy value from each of 35 slices in the $y - z$ plane). Hence forth we shall denote the 80 dimensional texture feature vector as Tex . Fig. II(e) shows the plot texture entropy values for two slices of diseased and normal patches.

Shape Asymmetry: We extend the concept of asymmetry (or anisotropy) to shape features. When healthy tissues are affected by progression of disease, it also affects spatial arrangement of voxels and hence their shape. Our aim is to exploit this irregularity for distinguishing between diseased and normal tissues. Shape features in the VOI are characterized by the 3D curvature of voxels. A detailed explanation of 3D curvature calculation is given in (<http://www.cs.ucl.ac.uk/staff/S.Arridge/teaching/ndsp/curvature3D.pdf>).

Shape asymmetry is calculated in a manner similar to texture anisotropy. The entropy of curvature values is determined from 9 sectors of each slice. If the curvature values have a wide distribution it indicates greater asymmetry in shape, leading to a higher entropy value. On the other hand low entropy values indicates less shape asymmetry. The shape asymmetry measure for a sector r is given by

$$Shape_{Asymmetry}^r = - \sum_{\theta} p_{\theta}^r \log p_{\theta}^r. \quad (2)$$

p_{θ}^r denotes the probability of distribution of curvature values in sector r , θ indicates the curvature values. Similar to texture anisotropy, the shape asymmetry measure is also a 80 dimensional feature vector. Henceforth the shape asymmetry vector is denoted as *Shape*. The above set of features give a 164 dimensional feature vector for a single 3D VOI.

Figs. 1(a) and (b) show a 2D patch around a diseased pixel, and the corresponding map of curvature values. Figs. 1(c) and (d) show a patch around a normal pixel and the corresponding curvature map. Figs. 1(e) and (f), respectively, show the plot of entropy values of texture and shape for 9 sectors of two slices from the VOI. The colormap of Figs. 1(b) and (d) are in the same range of [0, 5], with red regions denoting regions of high curvature. The curvature profile for the normal patch is quite regular as compared to the diseased patch. This is indicative of the fact that the curvature in diseased regions becomes distorted due to ulcerations or other abnormalities. Thus they lose the regularity observed in healthy tissues. This is corroborated by the plot in Fig. 1(f) where the diseased patches show higher entropy indicating greater randomness.

3 Experimental Results

3D T1-weighted spoiled gradient echo sequence (SPGE) images were acquired from 26 patients in supine position using a 3-T MR imaging unit (Intera, Philips Healthcare). The spatial resolution of the images was $1.02 \text{ mm} \times 1.02 \text{ mm} \times 2 \text{ mm}$, and the acquired volume dimension was $400 \times 400 \times 100$ voxels. The number of annotated voxels were 6827 from diseased regions, 5156 from normal and 3725 from background regions. An expert radiologist annotated each diseased region by outlining the region on each corresponding slice.

Features and Classifiers: We compare our method with two others, namely the Dual-Complex Tree Wavelet Transform (DTCWT) based method of [3] (*DCTWT*) and a shape asymmetry based method derived from the reflectional asymmetry measure of [12] (*Asy*). These two methods are recent developments in the field of disease and asymmetry detection. Although our method uses more features than these methods, it also highlights the challenges of localizing areas affected with Crohn’s diseases Each of the three methods was evaluated using three different classifiers, Random forests (RF), Support Vector Machines (SVM), and a Bayesian Classifier (BC). Random forests [5] have been successful in a variety of domains and compare favorably with other state-of-the-art

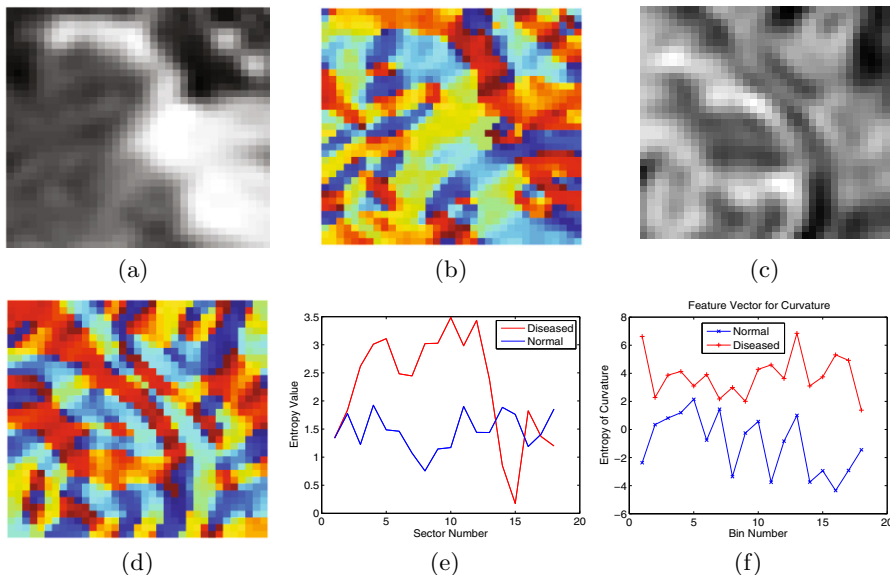


Fig. 1. (a)-(b) 2D patch around diseased voxel and corresponding curvature map; (c)-(d) normal patch and corresponding curvature map; plot of entropy values for (e) texture and (f) curvature of slices from VOI

algorithms [8]. A random forest is an ensemble of decision trees where each tree is trained with a different subset of the training data to improve the classifier's generalization ability. Training finds the set of tests that best separate the training data into different classes. Random forests and their variants have been used to detect abnormalities in mammograms [3], identify coronary artery stenoses [10] and semantic segmentation in CT images [14]. We use 50 trees for the RF classifier.

Support Vector Machines (SVMs) construct a hyperplane or set of hyperplanes in a high-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data of any class (so-called functional margin). In general larger the margin the lower the generalization error of the classifier. SVMs have also seen wide application in classification tasks like brain tumor segmentation [2][19], chest pathologies [1] and Glaucoma classification [7] among others. For SVMs we use the LIBSVM package [6] and define a radial basis function (RBF) as the kernel.

The default naive Bayesian classifier in MATLAB was the third classifier. A Bayesian classifier was chosen to highlight the linearly non-separable nature of the data, and the advantages of having a RBF kernel in SVMs. We have two classification stages for all classifiers. For all classifiers we employ 10-fold cross validation (leave-one-out with 10 subsets of the original data) approach.

Table 1. Quantitative measures for the *Stage 1* classification using different features and three classifiers. Values indicate mean and standard deviation. Sensitivity is the number of correctly detected intestine samples. Specificity is the number of correctly detected background samples. Accuracy is the total of number of correct detections (both background and intestine samples).

	Asy			DTCWT			Our Features		
	SVM	BC	RF	SVM	BC	RF	SVM	BC	RF
Accuracy (%)	80.4 ± 2.6	72.0 ± 2.3	79.9 ± 2.2	82.2 ± 2.4	71.3 ± 2.9	80.1 ± 2.5	88.1 ± 2.0	76.3 ± 6.4	86.8 ± 1.8
Specificity (%)	67.9 ± 1.8	41.5 ± 1.8	68.0 ± 1.7	68.1 ± 1.6	42.7 ± 1.7	67.6 ± 1.8	71.1 ± 1.8	54.1 ± 4.1	70.6 ± 2.2
Sensitivity (%)	86.2 ± 1.9	71.5 ± 1.4	84.6 ± 1.8	93.9 ± 2.7	78.3 ± 2.1	85.7 ± 1.9	97.9 ± 1.3	79.1 ± 4.5	96.1 ± 1.7

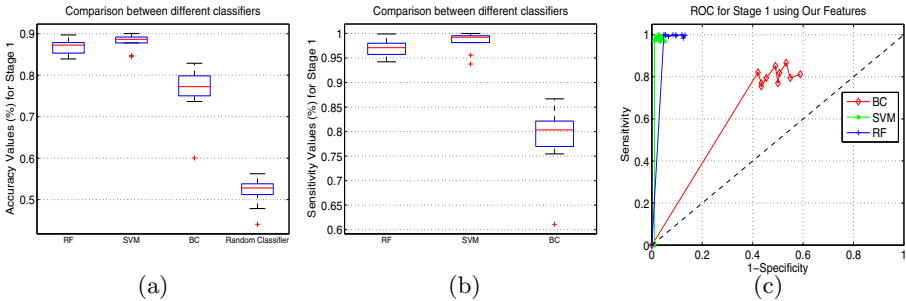


Fig. 2. Box plots for *Stage 1* classification: (a) Accuracy; (b) Sensitivity. (c) ROC curves for three classifiers using our features.

Classification Results for Stage 1. Table 1 shows different quantitative measures of the first classification stage. Here each sample is classified as either intestine or background. The highest classification accuracy is obtained using our features, the results of which are shown in the box plots of Fig. 2. In this stage we desire a high sensitivity or true positive rate (TPR) even at the expense of low overall accuracy. True positive refers to a intestine sample correctly classified as intestine. We do not want a intestine sample to be incorrectly labeled as background thus increasing the false negative rate (FNR). In such a situation, the diseased samples (which are part of intestine in the first stage) get classified as background and hence escape the scrutiny of the next stage. This is particularly undesirable in a clinical decision making system.

A comparison of receiver operating characteristics (ROC) curves of all three classifiers using our features for stage 1 is shown in Fig. 2 (c). RF and SVM give high sensitivity (more than 90%), but their specificity values are comparatively lower indicating a large number of false positives, i.e., many background samples are classified as intestine. The overall accuracy (i.e., correct classification percentage of both intestine and background samples) is lower than 86% in all cases. This again indicates a high number of false positives. This is not a

Table 2. Quantitative measures for individual and different combination of features using RF classifier

	<i>Int</i>	<i>Tex</i>	<i>Shape</i>	<i>Tex + Int</i>	<i>Shape + Int</i>	<i>Shape + Tex</i>
Accuracy (%)	77.1±2.3	81.6±2.1	79.1±2.7	79.2±1.3	79.5±2.4	82.3±1.3
Sensitivity (%)	79.3±3.2	86.9±2.1	82.3±1.9	83.1±3.1	83.8±2.3	86.6±2.8

disadvantage since these incorrectly labeled background samples are invariably identified as normal in *Stage 2* (as is evident from the results in Table 3 and Fig. 3). BC (which is good in classifying data with a linear decision boundary) has an overall accuracy less than 75%, which indicates the non-linearly separable nature of datasets. All classifiers perform better than a random classifier.

Importance of Different Features. Table 2 shows results of *Stage 1* classification using individual features and their combinations. The combination of texture and shape features performs closest to the values in Table 1. However, this does not diminish the contribution of intensity. A Student *t*-test on the values for *Tex + Shape* (Table 2) and *Our Features* (Table 1) gives $p < 0.032$ indicating statistically different results ($p < 0.05$ indicates that the two sets of results are statistically different, and hence significant). *t*-tests between *Tex* and *Tex – Int*, and *Shape* versus *Shape – Int* give $p < 0.04$, showing that inclusion of intensity statistics contributes to significant improvement.

DTCWT calculates a mean texture value across different orientations and scales while *Tex* calculates mean, variance, skewness and kurtosis across orientation and scales. Thus it is expected that *Tex* would be a more accurate measure than *DTCWT*, as reflected in Tables 2 and 1. *Asy* and *Shape* provide similar information as indicated by results of *t*-tests ($p = 0.13$).

Classification Results for Stage 2. Intestine samples from *Stage 1* are considered in *Stage 2* for further classification into diseased or normal. Table 3 shows different quantitative measures for *Stage 2* based on the *original* number of *diseased and normal* samples (not background samples) at the beginning of *Stage 1*. The values in Tables 1 and 3 are not directly related as Table 3 does not consider the background samples. In *Stage 2* true positive refers to the number of correctly classified diseased samples. BC’s accuracy and TPR is significantly worse compared to RF and SVM, thus reinforcing our inference about the non-linearly separable nature of the samples. RF and SVM, however, have similar performance with accuracies greater than 85% when using our features.

Fig. 3 shows the box plots of accuracy and sensitivity, and ROC curves when we consider *Stage 2* classification independently. Obviously the values will be higher than those reported in Table 3 (which are based on the original number of samples). The box plots and ROC curves indicate that a high percentage of each sample type is correctly classified by both SVM and RF. This is highly desirable because ultimately we would like to detect the diseased regions from abdominal MRI.

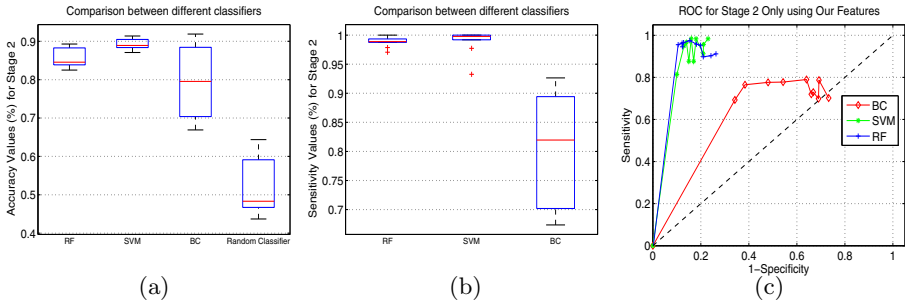


Fig. 3. Box plots for *Stage 2* classification: (a) Accuracy; (b) Sensitivity. (c) ROC curves for three methods using RF classifier.

Table 3. Quantitative measures for the second classification stage using different classifiers. Values indicate mean and standard deviation. Sensitivity is the number of correctly classified diseased samples. Specificity is the number of correctly classified normal samples. Accuracy is the total of number of correct classifications (both diseased and normal samples).

	Asy			DTCWT			Our Features		
	SVM	BC	RF	SVM	BC	RF	SVM	BC	RF
Accuracy (%)	81.2 ± 1.3	59.1 ± 0.9	81.3 ± 1.2	81.4 ± 1.4	58.4 ± 6.1	81.1 ± 1.2	87.5 ± 2.6	62.8 ± 5.4	87.2 ± 1.5
Specificity (%)	71.8 ± 3.1	36.1 ± 4.8	73.4 ± 2.4	72.8 ± 1.4	37.7 ± 2.7	72.2 ± 1.9	78.2 ± 1.7	42.3 ± 4.1	77.1 ± 1.6
Sensitivity (%)	84.5 ± 1.9	60.5 ± 1.2	84.9 ± 1.8	86.9 ± 1.7	61.3 ± 8.2	86.1 ± 1.9	91.9 ± 2.6	64.8 ± 10.7	90.4 ± 1.2

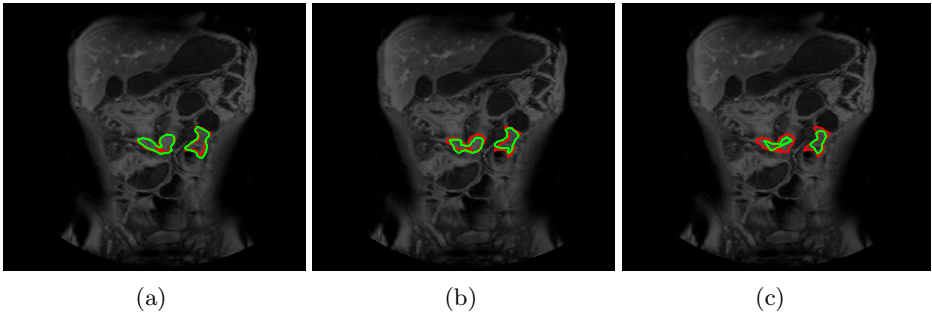


Fig. 4. Visual results for CD detection in Patient 16 with RF and different features: a) our features, b) *DTCWT* and c) *Asy*. Manually annotated diseased regions are shown in red while the result of automatic detection is shown in green.

Fig. 4 shows visual results for CD detection in Patient 16 using RF and the three features where each voxel is classified as diseased or normal. The manually annotated diseased regions are shown in red while the result of automatic detection is shown in green. The average Dice metric (DM) between manual

annotations and detected regions using our features is 86.5 ± 2.3 (RF), 77.2 ± 3.1 (BC) and 84.9 ± 3.2 (SVM). Our features do a good job in localizing the exact region afflicted with CD without explicit segmentation of the bowel wall.

4 Conclusion

We have proposed a method to identify VOI's in the abdominal MRI that are afflicted with Crohn's disease. Higher order intensity statistics, texture anisotropy and shape asymmetry are used to discriminate between diseased, normal and background regions. Higher order statistics capture image properties that are not discernible to the human eye. Our novel shape asymmetry measure is simple to compute than current approaches, and is informative in detecting diseased regions. Experimental results show that our designed feature vector performs better than *Asy* and *DTCWT*. Our results indicate that Crohn's disease can be detected from MR images, and thus reduce reliance on invasive procedures like colonoscopy and biopsy. With further improvements of our method in the future, we can hope to build a reliable detection and CD classification system.

Acknowledgements. This research was partly funded from the European Community's Seventh Framework Programme (FP7/2007-2013): the VIGOR++ Project (grant agreement nr. 270379).

References

1. Avni, U., Greenspan, H., Goldberger, J.: X-ray Categorization and Spatial Localization of Chest Pathologies. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 199–206. Springer, Heidelberg (2011)
2. Bauer, S., Nolte, L.-P., Reyes, M.: Fully Automatic Segmentation of Brain Tumor Images Using Support Vector Machine Classification in Combination with Hierarchical Conditional Random Field Regularization. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 354–361. Springer, Heidelberg (2011)
3. Berks, M., Chen, Z., Astley, S., Taylor, C.: Detecting and Classifying Linear Structures in Mammograms Using Random Forests. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 510–524. Springer, Heidelberg (2011)
4. Bhushan, M., Schnabel, J.A., Risser, L., Heinrich, M.P., Brady, J.M., Jenkinson, M.: Motion Correction and Parameter Estimation in dceMRI Sequences: Application to Colorectal Cancer. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part I. LNCS, vol. 6891, pp. 476–483. Springer, Heidelberg (2011)
5. Breiman, L.: Random forests. *Machine Learning* 45(1), 5–32 (2001)
6. Chang, C.-C., Lin, C.-J.: LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2, 27:1–27:27 (2011), <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
7. Cheng, J., Tao, D., Liu, J., Wong, D.W.K., Lee, B.H., Baskaran, M., Wong, T.Y., Aung, T.: Focal Biologically Inspired Feature for Glaucoma Type Classification. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 91–98. Springer, Heidelberg (2011)

8. Fuchs, T.J., Buhmann, J.M.: Computational pathology: Challenges and promises for tissue analysis. *Comp. Med. Imag. Graphics* 35(7-8), 515–530 (2011)
9. Julesz, B., Gilbert, E.N., Shepp, L.A., Frisch, H.L.: Inability of humans to discriminate between visual textures that agree in second-order statistics – revisited. *Perception* 2(4), 391–405 (1973)
10. Kelm, B.M., Mittal, S., Zheng, Y., Tsymbal, A., Bernhardt, D., Vega-Higuera, F., Zhou, S.K., Meer, P., Comaniciu, D.: Detection, Grading and Classification of Coronary Stenoses in Computed Tomography Angiography. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 25–32. Springer, Heidelberg (2011)
11. Kovalev, V.A., Petrou, M., Bondar, Y.S.: Texture anisotropy in 3D images. *IEEE Trans. Imag. Proc.* 8(3), 346–360 (1999)
12. Liu, Z., Smith, L., Sun, J., Smith, M., Warr, R.: Biological Indexes Based Reflective Asymmetry for Classifying Cutaneous Lesions. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 124–132. Springer, Heidelberg (2011)
13. Mary, J.Y., Modigliani, R.: Development and validation of an endoscopic index of the severity for crohn’s disease: a prospective multicentre study. *Gut*. 30(7), 983–989 (1989)
14. Montillo, A., Shotton, J., Winn, J., Iglesias, J.E., Metaxas, D., Criminisi, A.: Entangled decision forests and their application for semantic segmentation of ct images. In: MICCAI, pp. 184–196 (2011)
15. Pauly, O., Glocker, B., Criminisi, A., Mateus, D., Möller, A.M., Nekolla, S., Navab, N.: Fast Multiple Organ Detection and Localization in Whole-Body MR Dixon Sequences. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part III. LNCS, vol. 6893, pp. 239–247. Springer, Heidelberg (2011)
16. Petrou, M., Kovalev, V.A., Reichenbach, J.R.: Three-dimensional nonlinear invisible boundary detection. *IEEE Trans. Imag. Proc.* 15(10), 3020–3032 (2006)
17. Rimola, J., Rodriguez, S., Garcia Bosch, O., et al.: Magnetic resonance for assessment of disease activity and severity in ileocolonic Crohn’s disease. *Gut*. 58, 1113–1120 (2009)
18. Schunk, K.: Small bowel magnetic resonance imaging for inflammatory bowel disease (2002)
19. Verma, R., Zacharaki, E., Ou, Y., Cai, H., Chawla, S., Lee, S., Melhem, E., Wolf, R., Davatzikos, C.: Multiparametric tissue characterization of Brain Neoplasms and their recurrence using pattern classification of MR images. *Acad. Radiol.* 15(8), 966–977 (2008)
20. Zhan, Y., Shen, D.: Deformable segmentation of 3-d ultrasound prostate images using statistical texture matching method. *IEEE Trans. Med. Imag.* 25(3), 256–272 (2006)
21. Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D.: Multimodal classification of Alzheimer’s disease and mild cognitive impairment. *Neuroimage* 55(3), 856–867 (2011)

Out-of-Plane Motion Compensation in Cine-MRI

Mehmet Yigitsoy¹, Sonja Kirchhoff²,
Maximilian F. Reiser², and Nassir Navab¹

¹ Computer Aided Medical Procedures, Technische Universität München, Germany
yigitsoy@cs.tum.edu

² Institute of Clinical Radiology, Klinikum der Universität München, Germany

Abstract. Significant socio-economic burden of colonic motility disorders necessitates in-depth analysis of this pathology. Current analysis techniques are based on diameter measurements of colonic lumen on cine-MR images. Interleaved multi-plane acquisition makes it difficult to perform simultaneous measurements on the line of plane intersections due to the out-of-plane motion (OPM) caused by respiration affecting the underlying anatomy. Low temporal acquisition rate and dark-banding artifact are the challenging factors for OPM compensation. In this paper, we propose the use of manifold learning in combination with in-plane motion tracking for estimating OPM. We evaluate the effectiveness of our approach on 8 MR patient data sets. Experimental results show the good performance of our approach. The proposed method is independent of the acquisition rate and is not limited to this specific application.

Keywords: Motion compensation, manifold learning, motility, functional cine magnetic resonance imaging.

1 Introduction

Functional gastrointestinal disorders, such as diarrhea and chronic constipation, are conditions presenting with a significant socioeconomic burden. Chronic constipation is one of the most common of these conditions being one of the leading diagnoses for gastrointestinal (GI) disorders in the United States [13]. One important type of constipation is the slow-transit constipation associated with slow colonic transit time which is usually attributed to colonic motility disorders [1]. Therefore, it is necessary to study colonic motility in order to understand its effects on colon pathologies mentioned above leading to an improved and more adequate therapy in the end.

Existing well-established examination techniques such as manometry or scintigraphy are either invasive and inconvenient or expose patients to ionizing radiation creating the demand for a fast and non-invasive monitoring technique for the evaluation and quantification of colonic motility. On the other hand, functional cine magnetic resonance imaging (cine-MRI) allows for non-invasive, fast dynamic imaging with a superb soft tissue contrast [10].

The first use of cine-MRI for the analysis of colon motility was reported in [3] and the authors considered luminal diameter changes as motility parameter and manual diameter measurements were performed in the ascending, transverse and descending parts of the colon. Sets of 2D dynamic image sequences from the same imaging plane over time were acquired. Due to the respiratory gating, the sampling in time was irregular making the continuous tracking of lumen diameters infeasible. In [6], this approach was extended by addressing the irregular sampling in time and manual measurements on each 2D frame. A semi-automatic tool was developed to perform automatic diameter measurements on the data acquired during free breathing. However, measurements were still performed in 2D and the analysis of complex colon motion in 3D was not feasible by using this approach. Kutter et al. [9] proposed a multi-plane image acquisition of the colon by concentrating only on the descending part. In this setting, the images of the descending colon were acquired in multiple quasi-orthogonal planes in an interleaved way in order to perform simultaneous measurements on sagittal and coronal planes at various points on the line of intersection (LOI) between the planes and, then, to combine these values to have an approximation of the colon motion in 3D. However, since the acquisition in two planes is not simultaneous but sequential, the colon undergoes an out-of-plane movement caused by respiratory motion. This necessitates an OPM compensation in order to perform simultaneous diameter measurements at anatomically corresponding locations on two intersecting planes.

One approach to deal with this issue, is to approximate the breathing curve by fitting a sinusoidal to in-plane displacements via a Fourier analysis, as proposed in [9]. However, this idea is based on the assumption that the sampling rate per orientation plane is high enough according to the Nyquist-Shannon sampling theorem [12] for the recovery of the breathing curve with a frequency approximately 0.20 Hz. Apparently this is not feasible in our case where sampling rates per orientation are approximately 0.25 Hz. Another approach is the usage of the shared information, i.e. the intensity profiles (IP) along the lines of intersection (LOI), to have a rigid alignment as in [7]. However, due to dark banding artifact in the images caused by fast MRI acquisition [4], the IPs from two images present large variations resulting in uncorrelated intensity patterns along the LOI.

Therefore, in this work, we propose the use of manifold learning, a method that has recently been successfully applied to the image-based recovery of breathing motion [5], for the recovery and the compensation of the OPM in sagittal and coronal cine-MRI sequences. To this end, we perform dimensionality reductions on both sequences independently and use the average of the parameterizations of these embeddings to approximate the breathing curve in order to establish a basis for relating the sequences. In combination with in-plane motion tracking, significant compensation of the OPM can be achieved. The performance of the method is not affected by the data sampling rate as long as enough samples per orientation are collected.

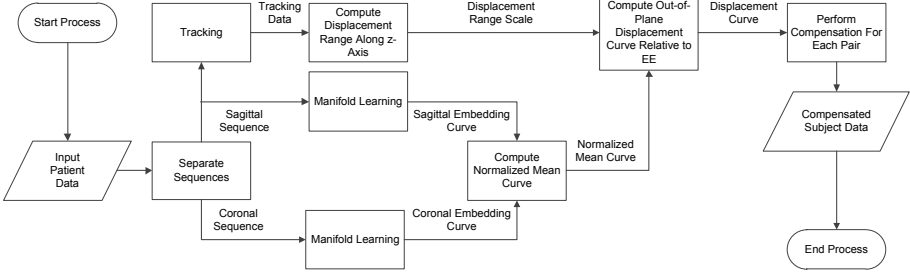


Fig. 1. Flow diagram of the OPM compensation algorithm

2 Materials and Methods

In this study 8 subjects were examined with functional cine-MRI [8]. The acquisition was performed using a half Fourier acquisition single shot turbo spin echo sequence (HASTE; TR: 476 ms, TE: 54 ms, SL: 5 mm, FOV: 320 x 400 mm, i-Pat-factor: 3.0, voxel size: 1.25 x 1.25 x 5.0 mm, duration: 8 min). Three orientations oblique to each other were set, examining 1 slice per orientation (transversal, sagittal, coronal) with a time interval of 4 s between image-sets to be able to physiologically image and track bowel motility. An alternating acquisition scheme, $\{\dots, \mathbf{u}_i^t, \mathbf{u}_{i+1}^s, \mathbf{u}_{i+2}^c, \mathbf{u}_{i+3}^t, \mathbf{u}_{i+4}^s, \mathbf{u}_{i+5}^c, \dots\}$, was used in order to equally sample data from each orientation. The transversal sequence is not used in this work as it is not used for colon motility analysis. For more details regarding the image acquisition protocol, please refer to [8].

2.1 Manifold Learning

The general idea of manifold learning is to project a manifold in high dimensional space \mathbb{R}^N to a low dimensional space \mathbb{R}^n , while preserving the local neighborhood. In our case, we consider one dimension of the ambient space for each image pixel, so N is corresponding to the resolution of the MRI images. For the low dimensional space, we set $n = 3$, in order to keep as much relevant information as possible while performing a significant dimensionality reduction. Considering k MR images $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ that are acquired over several breathing cycles in one orientation, the manifold learning \mathcal{M} assigns each image to a coordinate in the low dimensional space ϕ_i

$$\mathcal{M} : \mathbb{R}^N \rightarrow \mathbb{R}^n, \quad (1)$$

$$\mathbf{u}_i \mapsto \phi_i, \quad (2)$$

with $1 \leq i \leq k$. The suggestion that images lie on a low dimensional manifold in the ambient space seems to be justified because variations between neighboring slices are smooth, and furthermore, slices from the same breathing phase but

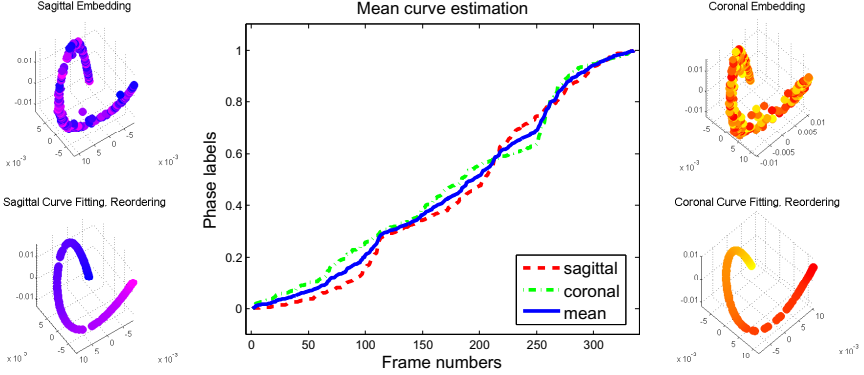


Fig. 2. Estimation of the mean motion curve via manifold learning. Curve fitting and parametrization are applied to each of the embeddings. In the figures on left and right, the color encodes the index before (upper) and after (lower) reordering.

different acquisition times share similar information. Moreover, since manifold learning techniques try to optimally preserve local information [2], meaning that similar images are mapped to similar positions in the low dimensional space, it is reasonable to use ϕ_i as an estimate for the respiratory phase.

We propose the application of Laplacian eigenmaps [2] for the respiratory phase estimation because the technique is well founded on mathematical concepts (Laplace Beltrami operator) and computationally efficient. Laplacian eigenmaps build upon the construction of a graph, which represents the neighborhood information of the data set. Subsequently, the graph Laplacian is applied to calculate a low-dimensional representation of the data that preserves the local neighborhood information in an optimal way.

We construct a graph with a node for each point \mathbf{u}_i and with edges connecting neighboring nodes. In order to deal with contrast differences between frames, we use cross correlation (NCC) as our similarity measure which is essential for neighborhood selection and weighting. We select for each image \mathbf{u}_i the l nearest neighbors, by evaluating the term $NCC(\mathbf{u}_i, \mathbf{u}_j)$. Further, heat kernel-based weights are assigned to the edges with

$$w_{ij} = e^{-(1-|NCC(\mathbf{u}_i, \mathbf{u}_j)|)^2 / (2 \cdot \sigma^2)} \quad (3)$$

and σ^2 the variance [2]. Once the neighborhood graph is constructed, the eigenvectors of the graph Laplacian provide the embedding map. After performing independent dimensionality reductions on coronal and sagittal sequences, these two embeddings need to be related in order to do further processing. The fact that both sequences are affected by the same breathing motion enables us to estimate the mean motion curve by using low dimensional embeddings of two sequences. Since the embeddings approximate curves in 3D, see Fig. 2, we first fit a 3rd order polynomial curve onto which we project original embedding points. Then, each curve is parametrized by point distances from one side of the curves.

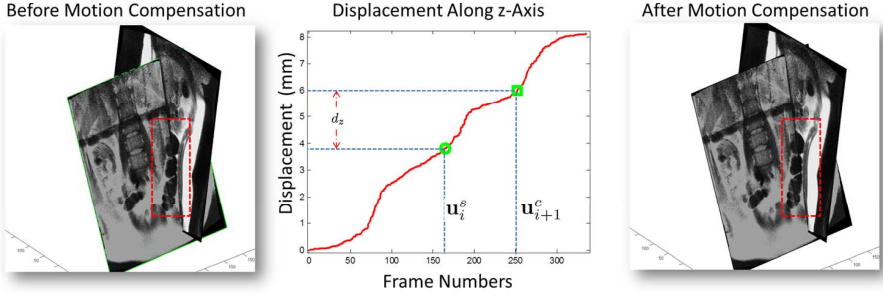


Fig. 3. Illustration of the motion compensation process. The region of interest around the LOI is highlighted where the difference after the compensation can be clearly observed by looking at the lumen borders. Contrast enhanced for better visibility.

The direction of parametrization is chosen so that both of them are parametrized from end-exhale (EE) to end-inhale (EI). After normalizing the parametrized curves, the mean breathing curve to be used as a basis for compensation is computed by taking the average of the normalized curves.

2.2 Out-of-Plane Motion Compensation

Since it is not possible to directly infer metric displacement values from the mean breathing curve, we need to compute the range of displacements in mm's, d_z , along the direction of dominant breathing motion, i.e. z-axis [11] in the reference coordinate frame. Then, we use this factor to scale the mean breathing curve in order to get the metric displacement values along the z-axis relative to the EE phase that is assumed to have no displacement.

In order to find d_z , we first estimate the range of in-plane displacements, d_p , in the sagittal orientation plane and then back project it onto the z-axis in order to approximate d_z . To this end, we perform tracking using block matching on the sagittal sequence to find a sequence of in-plane displacements, $\mathcal{D} = \{d_{p1}, \dots, d_{pk}\}$, relative to the first frame. A region that is highly influenced by the respiratory motion and close to the diaphragm is chosen for tracking. This is the visible part of the liver in our case. In-plane displacement range, d_p , is computed as $d_p = \max(\mathcal{D}) - \min(\mathcal{D})$. This value is projected on z-axis as

$$d_z = \begin{cases} \frac{d_p}{\sin(\alpha)}, & \text{if } \alpha > 0 \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where α is the angle between the sagittal plane normal and the z-axis calculated from the DICOM orientation data. The sagittal plane is chosen for estimating d_z since it is more aligned with the z-axis than the coronal plane. Once d_z is obtained, the mean breathing curve is updated by scaling it with d_z . Then, for each pair $\{\mathbf{u}_i^s, \mathbf{u}_{i+1}^c\}$, we compute the difference between their z-displacements relative to the EE phase, see Fig. 3. We shift the coronal plane by applying a

translational transform along the z-axis to bring it to the breathing phase of the sagittal plane leading to real anatomical correspondence along the LOIs.

3 Experiments and Results

We have conducted experiments on 8 different patient data to evaluate the performance of the proposed method. For each data set, we first separated coronal

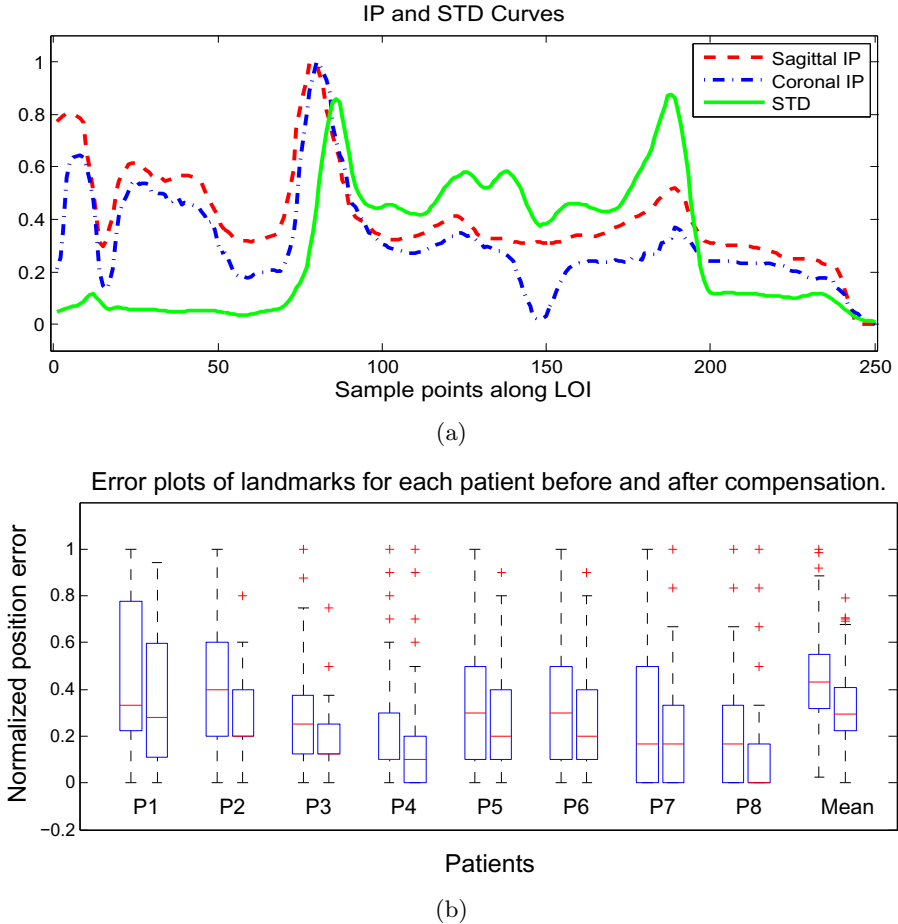


Fig. 4. (a) Mean IPs for each orientation plane and the intensity STD curve along the LOI. The landmark is searched for in the vicinity of the first peak from the right side in the STD curve. (b) Plot of normalized landmark errors for each patient. Each patient data has 336 pairs of coronal and sagittal to be aligned. Therefore, in this figure, each box pair represents error statistics for 336 image pairs before (left) and after (right) compensation.

and sagittal 2D sequences each consisting of 336 frames with a size of 256x320 pixels. Then the proposed method was applied to compensate for the OPM. Sagittal sequences of each subject underwent a tracking in order to estimate the in-plane-displacement range. In each case, Laplacian eigenmaps was used as the manifold learning technique. We set neighborhood size to 70 which empirically seemed to be sufficient for our application.

As discussed in Section 4, direct use of IPs for alignment is not possible due to large variations along the LOIs. Due to the same reason, we can not use IPs directly for validation purposes. Instead, we used a landmark that is affected by the breathing motion and that can easily be located in both of the IPs before and after the application of the motion compensation technique. A good landmark candidate is the edge of the liver which lies on the LOI and has a higher contrast that can be distinguished from the rest. In the case of a simultaneous acquisition, the positions of this landmark would match on both planes. Therefore, we used the difference between the positions of this landmark to assess the performance of compensation. A reduction of the difference indicates a good performance of the compensation technique. The location of the landmark in IPs was detected by searching for the maximum intensity around a predefined point. This point is associated with a high intensity standard deviation (STD) in time, see Fig. 4(a). There are also other points having high STD values, but, this one can be simply identified by picking the first one from the top.

For each subject, we performed difference measurements on 336 different pairs. We plot the statistics of the landmark error for each patient in Fig. 4(b). In each case, compensation resulted in a reduction of landmark error showing the effectiveness of the proposed approach.

4 Discussion and Conclusion

For an accurate analysis of the colon motility, motion due to respiration must be suppressed both in- and out-of-plane. It is of great importance to compensate for the OPM since the subsequent diameter measurements are directly based on LOIs. Therefore, in this work, we proposed a novel OPM compensation technique for dynamic, multi-plane, cine-MRI sequences of the colon. Experimental results show the good performance of our method and high potential for being used in similar scenarios with multi-plane acquisitions and OPM artifacts.

The novelty of our method is based on the use of manifold learning for the estimation of out-of-plane breathing motion on multiple planes. This gives a ground for relating sagittal and coronal planes in terms of breathing motion. Since low dimensional embedding gives a relative distribution of high dimensional data in the low dimensional space, it is not possible to deduce metric displacements using manifold learning. Therefore, in-plane motion tracking on sagittal sequences is included to assign metric values to the low dimensional embeddings relative to EE phase. Correlations between tracking and manifold

learning for sagittal sequences were above 0.95 in average. This also reveals the effectiveness of manifold learning in detecting motion patterns.

Due to the nature of the problem, a simple validation approach based on IPs was not feasible. Dark banding artifact along the LOIs did not allow us to use IPs directly for the evaluation. Instead, we detected landmarks along LOIs that are highly influenced by the breathing motion. By measuring the difference between positions of landmarks on two planes before and after motion compensation, we were able to evaluate the performance of the proposed method. Experimental results revealed the possibility of recovering the out-of-plane breathing motion using a combination of in-plane motion tracking and manifold learning on image sequences.

Acknowledgements. This work was funded by DFG (German Research Foundation).

References

1. Bassotti, G., de Roberto, G., Castellani, D., Sediari, L., Morelli, A.: Normal aspects of colorectal motility and abnormalities in slow transit constipation. *World J. Gastroenterol.* 11(18), 2691–2696 (2005)
2. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation* 15(6), 1373–1396 (2003)
3. Buhmann, S., Kirchhoff, C., Wielage, C., Mussack, T., Reiser, M.F., Lienemann, A.: Assessment of large bowel motility by cine magnetic resonance imaging using two different prokinetic agents: a feasibility study. *Invest. Radiol.* 40(11), 689–694 (2005)
4. Derakhshan, J.J., Griswold, M.A., Nour, S.G., Sunshine, J.L., Duerk, J.L.: Characterization and reduction of saturation banding in multiplanar coherent and incoherent steady-state imaging. *Magn. Reson. Med.* 63(5), 1415–1421 (2010)
5. Georg, M., Souvenir, R., Hope, A., Pless, R.: Manifold learning for 4d ct reconstruction of the lung. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2008*, pp. 1–8. IEEE (2008)
6. Glocker, B., Buhmann, S., Kirchhoff, C., Mussack, T., Reiser, M., Navab, N.: Towards a computer aided diagnosis system for colon motility dysfunctions. In: *SPIE Medical Imaging, San Diego, California, USA (February 2007)*
7. Kim, K., Habas, P.A., Rousseau, F., Glenn, O.A., Barkovich, A.J., Studholme, C.: Intersection based motion correction of multislice mri for 3-d in utero fetal brain image formation. *IEEE Trans. Med. Imaging* 29(1), 146–158 (2010)
8. Kirchhoff, S., Nicolaus, M., Schirra, J., Reiser, M.F., Göke, B., Lienemann, A.: Assessment of colon motility using simultaneous manometric and functional cine-mri analysis: preliminary results. *Abdominal Imaging* 36(1), 24–30 (2011)
9. Kutter, O., Kirchhoff, S., Berkovic, M., Reiser, M., Navab, N.: Spatio-temporal registration in multiplane mri acquisitions for 3d colon motility analysis. In: *SPIE Medical Imaging, San Diego, California, USA (February 2008)*
10. Lienemann, A., Sprenger, D., Steitz, H.O., Korell, M., Reiser, M.: Detection and mapping of intraabdominal adhesions by using functional cine mr imaging: preliminary results. *Radiology* 217(2), 421–425 (2000)

11. Rohlfing, T., Maurer Jr., C.R., O'Dell, W.G., Zhong, J.: Modeling liver motion and deformation during the respiratory cycle using intensity-based nonrigid registration of gated MR images. *Medical Physics* 31, 427 (2004)
12. Shannon, C.E.: Communication in the presence of noise. In: *Proceedings of the IRE*, vol. 37(1), pp. 10–21 (1949)
13. Talley, N.J.: Functional gastrointestinal disorders as a public health problem. *Neurogastroenterology & Motility* 20(suppl. 1), 121–129 (2008)

Real-Time Phase Boundary Detection for Colonoscopy Videos Using Motion Vector Templates

Ruwan Nawarathna¹, JungHwan Oh¹, Jayantha Muthukudage¹,
Wallapak Tavanapong², Johnny Wong², and Piet C. de Groen³

¹ Department of Computer Science and Engineering, University of North Texas,
1155 Union Circle #311366, Denton, Texas 76203-5017, USA
{rdn0025, Junghwan.Oh, mjk0129}@unt.edu

² Department of Computer Science, 226 Atanasoff Hall, Iowa State University,
Ames, Iowa 50011-1040, USA
{tavanapo, wong}@cs.iastate.edu

³ Department of Medicine, Division of Gastroenterology and Hepatology,
Mayo Clinic, 200 First St. S.W., Rochester, Minnesota 55905, USA

Abstract. Colonoscopy is the preferred screening method currently available for detection of colorectal cancer and its precursor lesions, colorectal polyps. However, recent data suggest that there is a significant miss rate for the detection of polyps in the colon during colonoscopy. Therefore, techniques for real-time quality measurement and feedback are necessary to aid the endoscopist towards optimal inspection to improve the overall quality of colonoscopy during the procedure. A typical colonoscopy procedure consists of two phases: an insertion phase and a withdrawal phase. One of the most essential tasks in real-time fully automated quality measurement is to find the location of the boundary between insertion and withdrawal phases. In this paper, we present a method based on motion vector templates to detect the phase boundary in real-time. The proposed method detects the phase boundary with a better accuracy and a faster speed compared to our previous method.

Keywords: Colonoscopy, phase boundary, end of insertion, motion vectors, camera motion estimation, and motion vector templates.

1 Introduction

Colonoscopy is the preferred screening modality for prevention of colorectal cancer---the second leading cause of cancer-related deaths in the US [1]. A typical colonoscopy procedure consists of two phases: an insertion phase and a withdrawal phase. The main purpose of the insertion phase is to reach the end of the colon, whereas in the withdrawal phase, careful inspection of all visible mucosa, tissue sampling, polyp removal, etc., are performed. Despite being the preferred screening modality, recent data suggest that there is a significant miss-rate in the detection of even large polyps during colonoscopy [2]. The miss-rate may be related to the experience of the endoscopist and the location of the lesion in the colon, but no prospective studies related to this have been done thus far. The American Society for Gastrointestinal Endoscopy

has suggested many guidelines for best practices in colonoscopy as described in [2] which includes the duration of the withdrawal phase, the average polyp detection rate, and the thorough of inspection of the colon mucosa. In [3], six quality metrics are proposed which are based on the durations of the insertion and the withdrawal phases. Therefore, accurate detection of the phase boundary (end of insertion (*EOI*)) between the insertion phase and the withdrawal phase is very essential in fully automated quality analysis of colonoscopy procedures.

The best way to detect the *EOI* is to analyze the motion of the colonoscopy camera, specifically, the z-directional motion (i.e., dolling camera motion (*DCM*)). The reason is that the colonoscope moves in forward and backward directions inside the colon during colonoscopy [2]. An accurate estimation of the camera motion can be obtained by analyzing the change that occurred between two consecutive images (frames) (i.e., a frame pair) in the video due to the movement of the camera. This change can be represented by using motion vectors. A motion vector represents the displacement of an area (usually a macroblock) that occurred due to the movement of the camera. The major challenges in accurately detecting the *EOI* with this approach are (1) colonoscopy frames have various artifacts such as out of focus (i.e., blurriness), specular reflection, stool, and water which can make the motion vector generation process imprecise, (2) traditional camera motion models such as the affine model [4] produce inaccurate motions often for colonoscopy, and (3) the method must complete all tasks in real-time. Specifically, our colonoscopy videos output 30 frames per second in MPEG-2 format. So, in order to achieve real-time processing, motion estimation of a frame pair must be completed within 66 milliseconds (ms) (i.e., 33ms per frame x 2).

In this paper, we propose a new method to detect the *EOI* based on motion vector templates. This method attempts to analyze the motion vector distribution in the four corners of the frame pair in order to predict the camera motion. So, proposed method does not depend on heavy computations as in traditional models such as the affine model [4]. The motion vectors are generated using an optical flow block matching algorithm [5-7] which is more suitable for the motion vector generation in colonoscopy frames. Motion vectors are obtained only on the frames that guarantee to provide accurate motions. Therefore, the primary contributions of this paper are (1) we propose a new algorithm to estimate camera motions in colonoscopy videos more accurately using motion vector templates and (2) our new algorithm offer a very large improvement in the speed which leads to significantly better real-time *EOI* detection of colonoscopy videos compared to our previous method [4].

The remainder of this paper is organized as follows. Related work in the field of colonoscopy frame processing and a brief analysis of our previous work [4] are presented in Section 2. The proposed phase boundary detection technique is described in Section 3. In Section 4, we discuss our experimental setup and results. Finally, Section 5 presents some concluding remarks.

2 Related Work

The related works on colonoscopy can be divided into three main categories: (1) processing of frames for tasks such as non-informative frame detection, stool frame

detection, and many others [8-10], (2) detection of abnormalities such as colorectal polyps [10], and (3) analysis of the quality of colonoscopy procedures [3].

The only work on real-time *EOI* detection that can be found in the literature is our previous work outlined in [4]. In that work the *EOI* is detected by applying a three step approach; (1) motion vector generation using color-based block matching, (2) camera motion estimation using the affine model, and (3) accumulation of *DCM* values. Due to the following issues, the accuracy of the *EOI* detection in our previous work is not satisfactory. Color-based block matching for motion vector generation is heavily dependent on the color information of the images. But, the colonoscopy frames have a limited color range [1, 2]. Also, since colonoscopy frames have a variety of artifacts such as blurriness, stool, and water, color-based block-matching method generates many flawed motion vectors. Moreover, the affine camera model [4] is very sensitive to outliers and generates incorrect *DCM* values when diverse motion vectors are present. Due to these incorrect *DCM* values, unnecessary local maxima (i.e., peaks) will be generated during *DCM* accumulation; consequently an incorrect point is detected as the phase boundary. The proposed method overcomes these issues and detects the *EOI* with 22% better accuracy and with 40-times better speed when compared to the previous work as described in Section 3.

3 Proposed Method

The proposed method has four main steps: (1) preprocessing of colonoscopy frames to discard/enhance unsuitable frames for motion vector generation, (2) motion vector generation using optical flow, (3) camera motion estimation using motion vector templates, and (4) detection of the phase boundary by analyzing cumulative *DCMs*. In the following sections each step will be discussed in detail.

3.1 Frame Preprocessing

As mentioned in Section 1 various artifacts such as blurriness, specular reflection, stool, water, therapeutic instruments, etc in colonoscopy frames can create errors in the motion vector generation process. Hence, preprocessing of frames is mandatory to get accurate motion vectors. If at least one frame of a frame pair in a video stream is a blurry frame then that frame pair is ignored from the motion vector generation process using the method outlined in [8]. Stools can be found in most sections in the colon. By their nature, these stools can float inside the colon and generate object motions which can combine with the camera motions. This can provide very unreliable motion vectors. Therefore, we eliminate the frame pairs having at least one frame with a stool percentage greater than a certain threshold (set to 50% based on experiments) using the technique proposed in [9]. In addition to these frame pairs, we discard frame pairs if the frames in the frame pair are not sufficiently correlated (see Fig. 1(a)). We calculate a correlation score between the two frames in each frame pair as expressed in equation (1). Then, only the frame pairs having correlation scores within a threshold value range are selected for motion vector generation. We set the threshold range to 0.89 - 0.99. An upper threshold is used to remove highly correlated frames as such frame pairs produce very few motion vectors since they are very similar.

$$\text{Corr}(A, B) = \frac{E[(I_A - \mu_A)(I_B - \mu_B)]}{\sigma_A \sigma_B}. \quad (1)$$

In equation (1), I_A and I_B are the intensities; μ_A and μ_B are the mean intensities; and σ_A and σ_B are the standard deviations of intensities of consecutive frames A and B, respectively. E is the expected intensity value. Moreover, colonoscopy frames have significant amount of noise. Also, in some frames, the brightness is inconsistent across the frame and in some frames the contrast is low. The noise is reduced by using a Gaussian filter of size 5x5 [7], and the brightness distribution is made to be uniform by applying histogram equalization [7]. Fig. 1(a) and 1(b) show a discarded and a retained frame pair for motion vector generation.

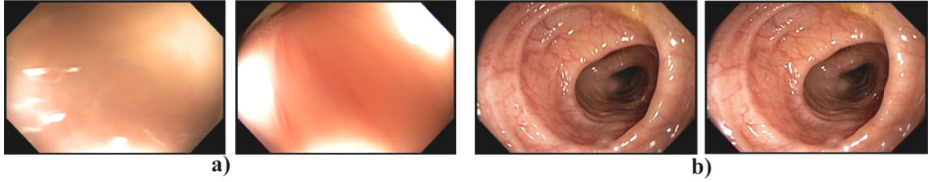


Fig. 1. a) A discarded frame pair because they are not sufficiently correlated, and b) a retained frame pair for motion vector generation

3.2 Motion Vector Generation Using Optical Flow Block Matching

We generate motions vectors on every 10th frame pair in the video stream which pass the preprocessing stage. This is done because camera motion of 10 consecutive frames is very similar. Color-based block matching [4] and the optical flow [5-7] are the most widely used motion vector generation algorithms. Due to limited color information in colonoscopy frames [10] the accuracy of color-based block matching is not sufficient for motion vector generation. Numerous methods have been proposed to compute the optical flow from adjacent frames in a video. We tried a sparse optical flow method (iterative Lucas-Kanade method with pyramids [5]), a dense optical method (Horn and Schunck algorithm [6]) and a block matching based optical flow method [7]. Since colonoscopy frames lack distinctive features, the Lucas-Kanade method tends to select stool and specular reflection areas as feature points. Because stool and specular reflection areas produce object motions, most of the generated motion vectors by this method are related to object motions. The Horn-Schunck method produces a dense optical flow [6]. This method is computationally very expensive and it also requires a very good motion vector filtering algorithm since it can generate lots of dissimilar motion vectors. Due to these reasons iterative Lucas-Kanade and Horn-Schunck methods are not suitable solutions to our problem.

Optical flow block matching is known to be more precise than the global Horn-Schunck method, and may also be faster since no iterative scheme is needed. Also, this method ensures that the effect of object motions (created by stools and specular reflection) on motion vectors is minimized considerably. Thus, we decided to use the optical flow block matching algorithm to generate motion vectors. This method

attempts to divide both previous and current frames into blocks, and then computes the motions of these blocks using optical flow [5, 7]. For each $m \times m$ block (B_k) centered around pixel (x,y) in frame k , we obtain a search area S in frame $k-1$ with B_k at the center block. The size of the search area S is $(m+2p) \times (m+2p)$ where p indicates the search range in pixels. Then, we compute the sum of square differences (SSD) between B_k and all possible $m \times m$ blocks in S as given in the equation (2). The $m \times m$ block (B_k') in S centered around (x',y') , which gives the lowest SSD is selected as the matching block. The displacement vector given by $u = x - x'$; $v = y - y'$ is the motion vector between B_k and B_k' . $f_k(x,y)$ is the intensity of the pixel at (x,y) . We experimentally found that a block size of 8×8 pixels, a search area size of 16×16 pixels and a SSD threshold of 128 are able to generate more accurate motion vectors for colonoscopy. Fig. 2(a) shows a typical result from motion vector generation. (The reason for having motion vectors only in four corner regions is explained in Section 3.3).

$$(u, v) = \underset{\substack{u=0,\dots,-p \\ v=0,\dots,p}}{\operatorname{argmin}} \sum_{i=0}^{-m-1} \sum_{j=0}^{m-1} (f_k(x+i, y+j) - f_{k-1}(x+i+u, y+j+v))^2. \quad (2)$$

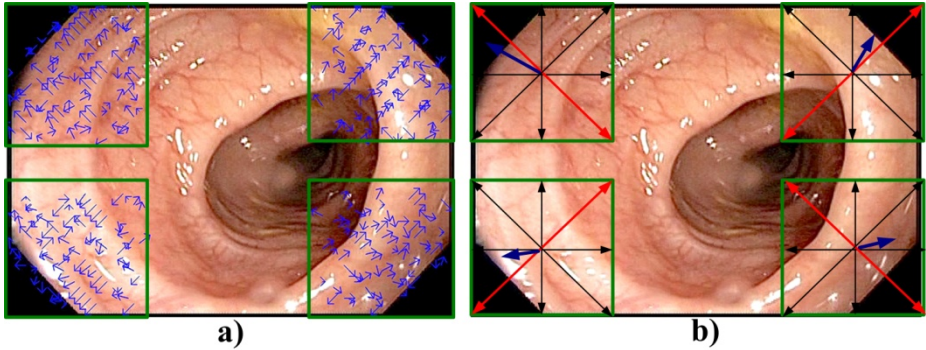


Fig. 2. a) Generated motion vectors from a frame pair in the four motion vector templates (arrow sizes are scaled up by a factor of 4) and b) their $TMVs$ (i.e., blue arrows; $TMVs$ are scaled so that they fit into their quadrants). Green squares represent the Template boundaries.

3.3 Camera Motion Estimation Using Motion Vector Templates

Motion vector templates have been studied previously in [11, 12]. A key aspect of our motion vector template method is the following: motion vectors in the four corners of our images show a unique pattern for each different type of camera motion. In the center part of the frame, we find either zero motion vectors (i.e., zero magnitude) (e.g. Z-directional translation) or motion vectors similar to four corners in the frame (e.g. X and Y motions). Hence, the camera motion can be estimated by analyzing the motion vector pattern in the four corners of a frame pair. Indeed, in the proposed method motion vectors are computed only in the four corners of a frame pair (see Fig. 2(a)). Four regions in the corners are called motion vector templates. Using our data set, we experimentally choose that the size of a Template as 13% of a frame, that is, 30% of

the width and 44% of the width of a frame (i.e., a Template is a square). Since, our aim is to find a specific pattern in the four motion vector templates; we represent the net motion in each Template region by one vector. This vector is computed by performing the vector addition of all motion vectors in each Template and then by calculating the mean vector as defined in equation (3). In a Template, the resultant vector is called its ‘‘Template Motion Vector (*TMV*)’’. In equation (3), mv_i^k is the k^{th} motion vector of the i^{th} Template and n_i is the number of motion vectors in the i^{th} Template. Four *TMVs* can be seen as illustrated in Fig. 2(b).

$$TMV_i = \frac{\sum_{k=1}^{n_i} mv_i^k}{n_i}. \quad (3)$$

Estimating the *DCM* by Computing the *DCM* Contribution. Forward and backward motions of the colonoscope can be estimated from positive *DCM* and negative *DCM*, respectively. Directions of the majority of motion vectors in a typical positive *DCM* are normally pointing from center to border. Hence, the directions of the four *TMVs* are also pointing from center to border (see Fig 3(a)). The directions of the majority of motion vectors and *TMVs* follow the opposite direction in a typical negative *DCM* (see Fig 3(b)). In general for a zooming motion, a similar behavior can be observed. Since, there is no zooming function available in a colonoscope; this pattern of *TMVs* can only be noticed when there is a *DCM*. Hence, we estimate the *DCM* when *TMVs* follow this pattern. We estimate the total *DCM* by calculating the *DCM* contribution from each Template. *DCM* contribution of a Template is calculated as the cosine of the angle that the *TMV* makes with the *DCM* axis multiplied by magnitude of the *TMV*. Use of cosine of the angle of *TMVs* ensures that more weight is assigned to the *TMVs* that are closer to the *DCM* axis (i.e., red line in each Template of Fig. 3). The rationale behind this design is that for a perfect *DCM* as seen in Fig. 3(a) and 3(b), *TMVs* are aligned with the *DCM* axis and the average of the magnitudes of *TMVs* represents the amount of *DCM*. For a non-perfect *DCM* as seen in Fig. 3(c), *TMVs* are away from the *DCM* axis and hence only a fraction of the magnitude represents the *DCM*. We compute the *DCM* contribution of a *TMV*, if it falls within the 45° range (in both directions, i.e., light blue and light green regions in Fig. 3) from the *DCM* axis. That means that a *TMV* within the 45° range from the *DCM* axis provides information about movement along the colon axis. *TMVs* which fall outside the 45° range carry essentially no information related to *DCM*. For this reason we ignore *TMVs* which fall outside the 45° region when calculating *DCM*.

The total *DCM* estimation process can be formulated as follows. We represent the magnitude and the angle of a k^{th} *TMV* which provides a non-zero positive *DCM* contribution as L_{pd}^k and θ_{pd}^k , respectively (see Fig. 3(a), *T1*, *T3* and *T4*). Also, L_{nd}^k and θ_{nd}^k represents the magnitude and the angle of the k^{th} *TMV* which provides a non-zero negative *DCM* contribution (see Fig. 3(c), *T2*). Here, *pd* stands for positive dolling and *nd* stands for negative dolling. In Fig. 3(c), *T1* has a positive *DCM* contribution and its value can be calculated as $L_{pd}^1 \cos \theta_{pd}^1$ and *T2* has a negative *DCM* contribution and its value can be calculated as $-L_{nd}^1 \cos \theta_{nd}^1$ and so on. We calculate the total *DCM* of a frame pair by taking *DCM* support count into consideration. The positive *DCM* support count (*PDSC*) is defined as the number of non-zero positive *DCM*

contributions and the negative *DCM* support count (*NDSC*) is defined as the number of non-zero negative *DCM* contributions. To be considered as a *DCM*, one of *PDSC* or *NDSC* must be at least two (experimentally decided). If we encounter equal (two) *PDSC* and *NDSC*, then we assign a zero *DCM*. Based on this model, we compute the average of positive *DCM* contributions for positive *DCM*, and the average of the negative *DCM* contributions for negative *DCM* as defined in equation (4).

$$DCM = \begin{cases} \frac{\sum_{k=1}^{PDSC} L_{pd}^k \cos \theta_{pd}^k}{PDSC}, & PDSC > NDSC \text{ and } PDSC \geq 2, \\ -\frac{\sum_{k=1}^{NDSC} L_{nd}^k \cos \theta_{nd}^k}{NDSC}, & NDSC > PDSC \text{ and } NDSC \geq 2, \\ 0, & \text{Otherwise.} \end{cases} \quad (4)$$

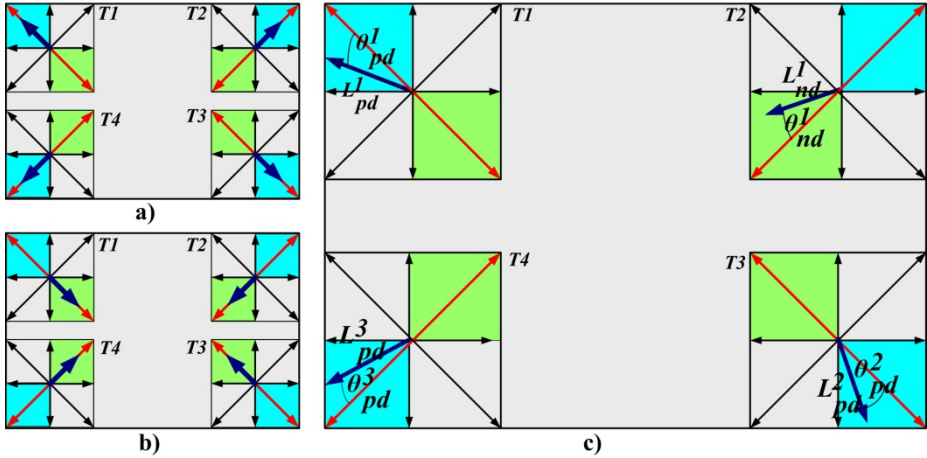


Fig. 3. An example of a) a perfect positive *DCM*, b) a perfect negative *DCM* and c) a *DCM* contribution calculation. Since $PDSC > NDSC$, (c) represents a positive *DCM* and it is also an example for a non-perfect positive *DCM*. In a), b), and c), *T1-T4* are four templates, red lines represent the *DCM* axes, light blue and light green regions represent the positive *DCM* and negative *DCM* supporting regions, respectively. The dark blue arrows show *TMVs*.

3.4 Estimation of the Phase Boundary/End of Insertion (*EOI*)

The estimation of the *EOI* is done by analyzing the behavior of cumulative *DCM* (*CDCM*) values of selected frame pairs in the entire procedure in real-time. We keep track of the *CDCM* values to find local maxima (i.e., peaks) in which the peak value remains unchanged for at least 1 minute (experimentally decided) as shown in Fig. 4. The frame number corresponding to this peak is assigned as the current *EOI* (i.e., a candidate *EOI*). Later, if we encounter another candidate *EOI* which has a *CDCM* value that is greater than the *CDCM* value of the current *EOI*, then the current *EOI* is updated with that candidate *EOI*. The frame number of the most recent candidate *EOI*

will be chosen as the *EOI*, which is the phase boundary. Obviously, the last *EOI* will be the frame number that defines the maximum *CDCM* value of the entire procedure. *CDCM* values of a full colonoscopy stream can be plotted as shown in Fig. 4. Since the size of a Template in the new method is 13% of a frame, we process only 52% of the frame area for motion vector generation. Also, we use an optimized version of the optical flow block matching algorithm provided by OpenCV 2.0 [7]. In addition to those, we apply CPU multithreading for motion vector generation by dividing the total number of blocks among the available processors for parallel processing. The combination of the above three factors ensures that the proposed method performs the *DCM* calculation well within real-time.

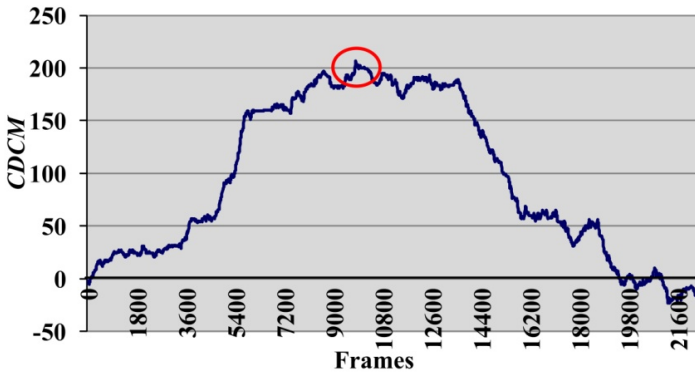


Fig. 4. A *CDCM* plot of a full colonoscopy video stream. The red oval is the detected *EOI*. Basically, this curve gives a snapshot of how the colonoscope moved through the colon in forward and backward directions during the procedure.

4 Experimental Setup and Results

The proposed method was implemented and integrated into SAPPAPHIRE [13] which is a framework developed for real-time capture and quality analysis of colonoscopy. We perform our experiments in a simulation mode where videos are used as real-time video streams. All experiments were done in a computer with Intel(R) Core i7 2600K, 64-bit, 3.40GHz processor and 8 GB memory. We conducted experiments on ordinary videos as well as on real colonoscopy video streams.

We created a video which contains five types of motions. The video was recorded at a valley outside of a building. Five different types of motions are: (1) slow forward (frame 0 to 611), (2) fast forward (frame 741 to 1151), (3) fast backward (frame 1161 to 1431), (4) slow backward (frame 1441 to 2281), and (5) zero motion (frame 611 to 731 and frame 1091 to 1161). In the *CDCM* plot shown in Fig. 5, the fast motions (2 and 3) can be seen as steep lines, and slow motions (1 and 4) can be seen as less-step lines. Also, zero motions (5) can be seen as straight lines. Therefore, our proposed method has accurately captured all motions that are contained in the video.

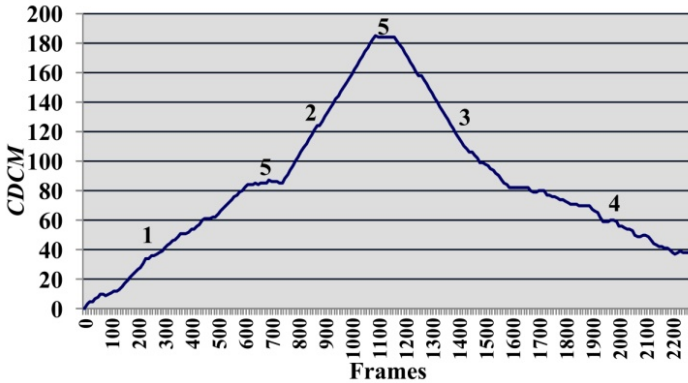


Fig. 5. CDCM plot of the ordinary video. Numbers 1-5 represent the five different types of motions present in the video.

Our real colonoscopy video set contains 146 videos. The average video length is 17.40 minutes and the frame size is 720 x 480. When the cecum is reached during colonoscopy, any remaining debris is removed followed by inspection of the appendiceal orifice, the ileocecal valve and if required or readily possible the distal terminal ileum. Clearing debris and inspection require about 1 to 2 minutes. Therefore, if the detected *EOI* falls within a 2 min range from the ground truth (i.e., before or after), then we consider it as a correctly detected *EOI* and vice versa. Outcomes from this experiment are summarized in Table 1. The results show that our new method detects the *EOI* with 82% accuracy. Therefore, compared to our previous method, we obtained a 22% improvement in accuracy. When comparing the average time difference with the ground truth of all videos of the two methods, the new method has shown a major improvement (81 seconds). Also, the proposed method provides a substantial gain (almost 40-times) in the average execution time per frame pair over the earlier method. The previous method satisfies the real-time constraint by a slight margin (by 13.70ms = 66 - 52.30). However, the new method satisfies the real-time constraint by a considerably larger margin (by 64.68 ms = 66 - 1.32).

Table 1. Effectiveness of the proposed phase boundary detection method

Description	Previous	New
Number of correctly detected videos	88	120
Accuracy of the <i>EOI</i> detection	60%	82%
Average time difference with the ground truth <i>EOI</i> (mm:ss)	02:54	01:33
Average execution time per frame pair (ms)	52.30	1.32

5 Conclusion

Finding the phase boundary is considered to be a very important task for fully automated quality analysis of colonoscopy. We have proposed and implemented a new algorithm for real-time phase boundary detection in colonoscopy using motion

vector templates. The proposed method detects the phase boundary with 82% accuracy which is 22% better than our previous algorithm. Also, the new algorithm is 40-times faster than our previous algorithm. Moreover, this method is easy to implement since it does not possess heavy computations compared to the previous method. Experimental results also demonstrate that this method works very well on estimating the dolling motions of ordinary videos. Accurate detection of the phase boundary leads to generation of many quality metrics, such as the ones proposed in [3]. A key problem affecting the accuracy of our current method is formed by frames containing remaining debris, air bubbles and water or specular reflections. In the future we will improve the accuracy by detecting and eliminating these frames.

Acknowledgments. This work is partially supported by NSF STTR-Grant No. 0740596, 0956847, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK DK083745).

References

1. American Cancer Society. Colorectal Cancer Facts & Figures (2008), http://www.cancer.org/docroot/STT/content/STT_1x_Cancer_Facts_and_Figures_2008.asp
2. Rex, D.K., Petrini, J.L., Baron, T.H., Chak, A., Cohen, J., Deal, S.E., Hoffman, B., Jacobson, B.C., Mergener, K., Pertersen, B., Safdi, M.A., Faigel, D.O., Pike, I.M.: Quality Indicators for Colonoscopy. *Gastrointestinal Endoscopy* 63, S16–S26 (2006)
3. Oh, J., Hwang, S., Cao, Y., Tavanapong, W., Liu, D., Wong, J., de Groen, P.C.: Measuring Objective Quality of Colonoscopy. *IEEE T. Bio-Med. Eng.* 56(9), 2190–2196 (2009)
4. Oh, J., Rajbal, M.A., Muthukudage, J.K., Tavanapong, W., Wong, J., de Groen, P.C.: Real-Time Phase Boundary Detection in Colonoscopy Videos. In: 6th International Symposium on Image and Signal Processing and Analysis, Salzburg, pp. 724–729 (2009)
5. Bouguet, J.Y.: Pyramidal Implementation of the Lucas Kanade Feature Tracker. Intel Corporation Microprocessor Research Labs 1(2), 1–9 (2001)
6. Horn, B., Schunck, B.: Determining Optical Flow. *Artif. Intell.* 17, 185–203 (1981)
7. Bradski, G., Kaehler, A.: *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media (2008)
8. Oh, J., Hwang, S., Lee, J., Tavanapong, W., Wong, J., de Groen, P.C.: Informative Frame Classification for Endoscopy Video. *Med. Image Anal.* 11(2), 110–127 (2007)
9. Muthukudage, J., Oh, J., Tavanapong, W., Wong, J., de Groen, P.C.: Color Based Stool Region Detection in Colonoscopy Videos for Quality Measurements. In: Ho, Y.-S. (ed.) *PSIVT 2011, Part I. LNCS*, vol. 7087, pp. 61–72. Springer, Heidelberg (2011)
10. Atasoy, S., Mateus, D., Lallemand, J., Meining, A., Yang, G.-Z., Navab, N.: Endoscopic Video Manifolds. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) *MICCAI 2010, Part II. LNCS*, vol. 6362, pp. 437–445. Springer, Heidelberg (2010)
11. Nguyen, T., Laurendeau, D., Branzan, A.: A Robust Method for Camera Motion Estimation in Movies Based on Optical Flow. *Int. J. Intell. Syst. Technol. Appl.* 9(3/4), 228–238 (2010)
12. Xiong, W., Lee, J.C.: Efficient Scene Change Detection and Camera Motion Annotation for Video Classification. *Computer Vision and Image Understanding* 71(2), 166–181 (1998)
13. Stanek, S.R., Tavanapong, W., Wong, J., Oh, J., Nawarathna, R., Muthukudage, J., de Groen, P.C.: SAPPHERE Middleware and Software Development Kit for Medical Video Analysis. In: 24th IEEE International Symposium on Computer-Based Medical Systems, Bristol, pp. 1–6 (2011)

Relaxed Conditional Statistical Shape Models and Their Application to Non-contrast Liver Segmentation

Sho Tomoshige¹, Elco Oost¹, Akinobu Shimizu¹, Hidefumi Watanabe¹,
Hidefumi Kobatake¹, and Shigeru Nawano²

¹ Graduate School of Bio-Applications and Systems Engineering,
Tokyo University of Agriculture and Technology,
Nakacho 2-24-16, Koganei-shi, Tokyo 184-8588, Japan
{c_r_oost, simiz}@cc.tuat.ac.jp

² Department of Radiology, International University of Health and Welfare,
Mita Hospital, 1-4-3 Mita, Minato-ku, Tokyo 108-8329, Japan

Abstract. This paper proposes a novel conditional statistical shape model (SSM) that allows a relaxed conditional term. The method is based on the selection formula and allows a seamless transition between the non-conditional SSM and the conventional conditional SSM. Unlike a conventional conditional SSM, the relaxed conditional SSM can take the reliability of the condition into account. Organ shapes estimated by the proposed SSM were used as shape priors for Graph Cut based segmentation. Results for liver shape estimation and subsequent liver segmentation show the benefit of the proposed model over conventional conditional SSMs.

Keywords: Conditional shape modeling, relaxation, liver segmentation.

1 Introduction

Graph Cut based segmentation [1] with a shape prior as regulating term in the optimization of the energy function has proven a valuable tool in medical image processing [2,3]. Using a set of image features that are extracted from the target image as the conditional term for a conditional SSM, for example as described in [4], a shape prior is estimated that will serve as a restricting term in the optimization of the energy function in Graph Cut segmentation. Given an appropriate shape prior the Graph Cut segmentation will improve. This benefit is the method's vulnerability as well: A poor quality shape prior will deteriorate the Graph Cut segmentation accuracy.

Relaxation of the condition, instead of applying the condition as a hard constraint, is essential in generating a shape prior. A relaxed conditional SSM should be able to bridge seamlessly between a non-conditional SSM and a conventional conditional SSM with hard constraints. Obtaining a method that allows such a seamless transition is the main aim of this paper, which is achieved by using the selection formula [5] for the calculation of the conditional covariance matrix and the conditional average.

Several conditional SSMs have been proposed. Baka et. al [6] propose a conditional SSM in which uncertainties of the conditions can be integrated. The algorithm

calculates a conditional covariance matrix, but does not calculate a conditional average. For a seamless transition between non-conditional SSM and conventional conditional SSM however, both the conditional covariance matrix and the conditional average are required. Syrkina et. al [7] propose a shape estimation method, that calculates a conditional distribution through a joint multivariate distribution of two statistical shape models; one representing the predictors and one for the shape that needs to be predicted. To minimize the prediction error, the number of shape modes retained for the two models is limited. In some cases this can lead to a considerable part of the training data that will be excluded from the model. Furthermore the algorithm requires an estimate of the noise variance, which is difficult to obtain.

The algorithm by de Bruijne et. al [4] uses the conventional conditional SSM with hard constraints, and extends it with ridge regression [8] to regularize the covariance matrix. Additionally, the introduction of ridge regression is an alternative approach to relax the conditional term of the SSM, because it allows the calculation of both the conditional covariance matrix and the conditional average. A ridge parameter of zero will result in the conventional conditional SSM with hard constraints, whereas a very large ridge parameter leads to the generic, non-conditional SSM. Hence, the range of the regulating term is between zero and infinity. The method proposed in this paper presents a more elegant transition, with a regulating term between zero and one. In addition, the identity matrix in ridge regression might not be suitable to relax the condition, because it enhances the influence of the weaker shape variations.

The benefit of the proposed method will be assessed by liver shape estimation, followed by liver segmentation in non-contrast CT images. Many state of the art algorithms for liver shape estimation and liver segmentation provide similar performance in relatively easy to segment images. The difficult to segment images, e.g. when the liver shape differs strongly from the average liver shape, remain challenging. Estimation of such shapes, based on image features, is difficult and can result in errors in the condition estimation. The subsequent segmentation is hampered by suboptimal shape estimation. This paper seeks the room for improvement for such cases. Presented results focus on difficult to segment images. Furthermore, because shape estimation by a ridge regression based conditional SSM is the closest related method, results of the proposed method will be compared with results obtained through ridge regression.

2 Conventional Conditional Statistical Shape Models

To train a level-set based SSM, a data set of N manually annotated images is used to create a signed distance map, in which voxel values represent the distance to the organ contour. Negative distances denote the organ's interior, positive distances signify the organ's exterior. The distance data is extracted to a one-dimensional column vector, sized M , and Principal Component Analysis is applied to create a SSM. Projection of the training samples onto the model results in the principal component score matrix b , as depicted in Fig. 1. Details on level set based SSM training can be found in [9].

To obtain conditional data, a number of features is calculated from the true label data for all training data samples. These features are combined in matrix X , which has N columns (number of training samples) and F rows (number of calculated features). Subsequently, an unseen test image is roughly segmented using maximum a posteriori

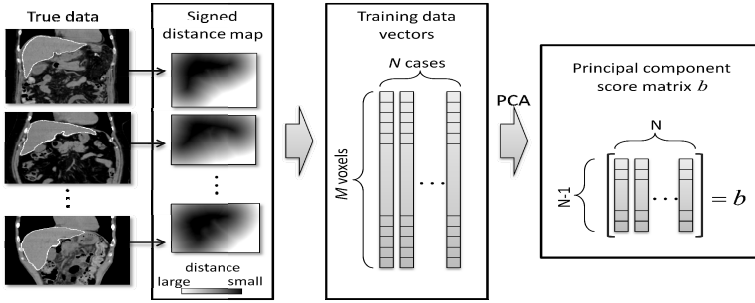


Fig. 1. Level set based statistical shape model training

estimation (MAP) [10], followed by the calculation of the same features that were obtained during the construction of matrix X . This set of features from the test image are combined in a column vector x_0 , and will serve as the condition for the conditional SSM. Using matrices b and X , and the calculated condition x_0 , the conditional average $\mu_{b|x_0}$ and the conditional covariance matrix $\Sigma_{bb|x_0}$ are defined as:

$$\mu_{b|x_0} = \mu_b + \Sigma_{bx} \Sigma_{xx}^{-1} (x_0 - \mu_x) \tag{1}$$

$$\Sigma_{bb|x_0} = \Sigma_{bb} - \Sigma_{bx} \Sigma_{xx}^{-1} \Sigma_{xb} \tag{2}$$

with μ_b signifying the regular shape model average, derived from matrix b and μ_x denoting the average set of conditional features, derived from matrix X . Σ_{xx} and Σ_{bb} are the covariance matrix of X and b respectively and Σ_{xb} and Σ_{bx} are mutual covariance matrices. Subsequently, the eigenvectors of the conditional covariance matrix $\Sigma_{bb|x_0}$ are rearranged in descending order of eigenvalues, after which the top L modes of variation are selected to form the conditional SSM space, as shown in Fig. 2.

3 Relaxed Conditional Statistical Shape Model

A drawback of the conventional conditional SSM is that the selected condition is considered to be reliable. Consequently, if the condition is inaccurate, the conditional SSM will be wrongly influenced and the reconstructed shape will be suboptimal.

To overcome this, a conditional SSM with a relaxed condition is proposed, which is constructed by using a selection formula [5]. To calculate conditional probability, equations (1) and (2) require the set of covariance matrices:

$$\begin{pmatrix} \Sigma_{xx} & \Sigma_{xb} \\ \Sigma_{bx} & \Sigma_{bb} \end{pmatrix}. \tag{3}$$

Following the selection formula, using only a limited range of the conditional features x results in a new covariance matrix V_{xx} , and equation (3) can be rewritten as:

$$\begin{pmatrix} V_{xx} & \Sigma_{xb} \\ \Sigma_{bx} \Sigma_{xx}^{-1} V_{xx} & \Sigma_{bb} - \Sigma_{bx} \left(\Sigma_{xx}^{-1} - \Sigma_{xx}^{-1} V_{xx} \Sigma_{xx}^{-1} \right) \Sigma_{xb} \end{pmatrix}. \tag{4}$$

The bottom right term in equation (4) signifies the conditional covariance matrix of b . Hence, given a conditional range around x_0 , resulting in a covariance matrix V_{xx} , the covariance matrix for b can be calculated by:

$$\Sigma_{bb|x_0} = \Sigma_{bb} - \Sigma_{bx} \left(\Sigma_{xx}^{-1} - \Sigma_{xx}^{-1} V_{xx} \Sigma_{xx}^{-1} \right) \Sigma_{xb}. \quad (5)$$

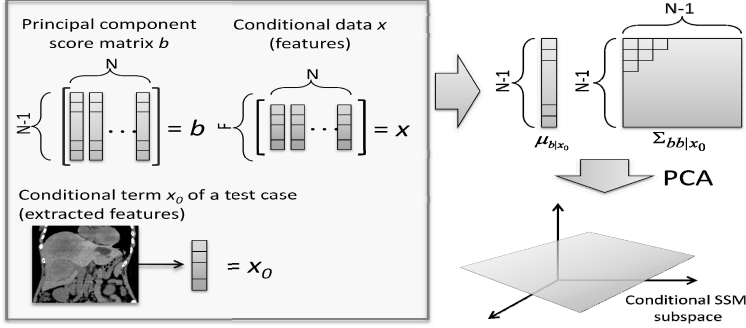


Fig. 2. Construction of the conditional statistical shape model

Note that if the full range of conditional features x is used, i.e. $V_{xx} = \Sigma_{xx}$, $\Sigma_{bb|x_0}$ equals Σ_{bb} , corresponding to a non-conditional SSM. In the other extremity, if the range of features x is limited, such that $x = x_0$, V_{xx} will become 0 and as a result equation (5) is identical to equation (2), representing the conventional conditional SSM. The average value for b , given the conditional range around x_0 becomes:

$$\mu_{b|x_0} = \mu_b + \Sigma_{bx} \left(\Sigma_{xx}^{-1} - \Sigma_{xx}^{-1} V_{xx} \Sigma_{xx}^{-1} \right) (x_0 - \mu_x). \quad (6)$$

Analyzing the extremities, similar behavior can be identified, as was observed for the covariance matrix of b . If the full range of conditional features x is used, i.e. $V_{xx} = \Sigma_{xx}$, $\mu_{b|x_0}$ equals μ_b (a statistical model without any conditions) and if the range of conditional features x is limited, such that $x = x_0$, V_{xx} becomes 0 and equation (6) will be identical to equation (1), representing the conventional conditional SSM.

The difference between the conventional and the relaxed conditional SSM is explained in Fig. 3. As visualized in Fig. 3a, the condition set in the conventional SSM results in a subspace that only comprises the area for which holds $x = x_0$. Contrary, for the relaxed conditional SSM, as shown in Fig. 3b, if a selection of samples for which holds that x approximates x_0 is taken, the fixed condition x_0 is replaced by a relaxed conditional range around x_0 , defined by covariance matrix V_{xx} .

The relaxed conditional SSM spans from a non-conditional SSM to a conventional conditional SSM. If a small range around x_0 is used to estimate parameter b (assuming a reliable condition), the model behaves like a conventional conditional SSM. If a large range around x_0 is used to estimate parameter b (assuming an unreliable condition), the model behaves like a non-conditional SSM. To bridge both SSMs, reliability parameters $\{\gamma_1, \gamma_2, \dots, \gamma_F\}$ ($0 \leq \gamma_i \leq 1$) are introduced. Since for every conditional feature the error might differ, the reliability parameter γ_i should be defined

individually. In order to reflect the covariance of conditions in the relaxation process, we multiply Σ_{xx} by the reliability parameters, simplifying V_{xx} to:

$$V_{xx} = \left((I - \Gamma)^{\frac{1}{2}} \right)^T \Sigma_{xx} (I - \Gamma)^{\frac{1}{2}} \quad (7)$$

with

$$\Gamma = \begin{pmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \gamma_F \end{pmatrix} \quad (8)$$

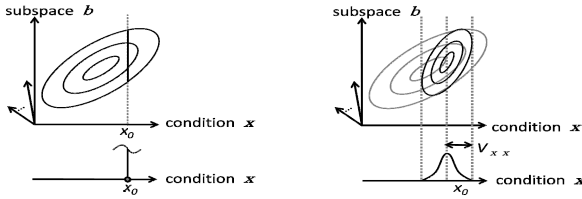


Fig. 3. Difference between conventional (left) and relaxed (right) conditional SSM

Substitution of equation (7) into (5) and (6) results in the final equations for the conditional covariance matrix and conditional the average in equations (9) and (10).

$$\Sigma_{bb|x_0} = \Sigma_{bb} - \Sigma_{bx} \left(\Sigma_{xx}^{-1} - \Sigma_{xx}^{-1} \left(\left((I - \Gamma)^{\frac{1}{2}} \right)^T \Sigma_{xx} (I - \Gamma)^{\frac{1}{2}} \right) \Sigma_{xx}^{-1} \right) \Sigma_{xb} \quad (9)$$

$$\mu_{b|x_0} = \mu_b + \Sigma_{bx} \left(\Sigma_{xx}^{-1} - \Sigma_{xx}^{-1} \left(\left((I - \Gamma)^{\frac{1}{2}} \right)^T \Sigma_{xx} (I - \Gamma)^{\frac{1}{2}} \right) \Sigma_{xx}^{-1} \right) (x_0 - \mu_x) \quad (10)$$

This differs from calculating the conditional covariance matrix and the conditional average through ridge regression, which is defined by:

$$\mu_{b|x_0} = \mu_b + \Sigma_{bx} (\Sigma_{xx} + \rho I)^{-1} (x_0 - \mu_x) \quad (11)$$

$$\Sigma_{bb|x_0} = \Sigma_{bb} - \Sigma_{bx} (\Sigma_{xx} + \rho I)^{-1} \Sigma_{xb} \quad (12)$$

in which ρ denotes the ridge parameter. In the proposed method the conditional covariance matrix and the conditional average are constructed by multiplication of the original covariance matrix with a value between zero and one, as expressed by equations (9) and (10). Ridge regression adds ρI to the original covariance matrix to calculate the conditional covariance matrix and conditional average. The benefit of the proposed method in deriving the conditional covariance matrix and conditional average is that the modifications are proportional to the variance of individual shape variations. Contrary, by adding a fixed ridge parameter, the weaker shape variations are

relatively strongly affected, whereas the influence on the stronger shape variations is limited.

It is worth mentioning that Σ_{xx} might be singular due to multi-colinearity of the features. In experiments, the number of samples was, compared to the number of conditional features, large enough to obtain non-singularity. In addition, from a large data base of features it is easy to select a set of features whose covariance matrix is not singular and which can still be effectively used as conditions for the SSM.

4 Estimation of a Shape Prior

The importance of a reliable conditional term in estimating the shape prior was stressed in section 1. To generate the shape prior, the following steps were performed:

1. Roughly extract the test image by maximum a posteriori estimation [10].
2. Project the MAP result onto the relaxed conditional SSM (described in Section 3) and define the parametric position as the search starting point.
3. Select the shape parameters 1 until $\lfloor L/3 \rfloor$, in which L denotes the number of shape parameters that represent 90% of the model's variation.
4. Using Powell's method [11], with the Jaccard Index as objective function, optimize the shape parameters for the projected MAP result.

To avoid local minima, the optimization is done in three subsequent steps: first for shape parameters 1 until $\lfloor L/3 \rfloor$, then for 1 until $\lfloor 2L/3 \rfloor$ and finally for 1 until L .

5 Experimental Setup and Results

The total data set consisted of 144 non-contrast abdominal CT images. The image size was $512 \times 512 \times 154 \sim 807$ voxels with a resolution of $0.546 \sim 1.00$ mm/voxel. The data was subsampled by a factor 2. The first 48 cases were used to train the SSM, the second 48 cases were used to decide and evaluate the Graph Cut parameters, to optimize the reliability parameter γ and to optimize ρ . The third 48 cases, that were available for testing, were divided into two categories: easy to segment and difficult to segment. The state of the art methods in liver segmentation all show acceptable segmentation results for easy cases. For difficult cases however improvements in segmentation can be achieved. This paper therefore will mainly focus on the set of 24 difficult cases. The 24 easy cases will only be discussed briefly. To distinguish between easy and difficult cases, the shape estimates for the 48 test cases were created using a standard, non-conditional SSM. The 24 cases that showed the lowest Jaccard Index after subsequent Graph Cut segmentation were marked as difficult cases.

A set of 20 features, derived from manual labels (training) or from MAP results (testing), was used as conditional terms. Among the features were the object length in x , y and z direction, surface areas of the projected object on sagittal, coronal and axial planes, the object's volume and histogram derived parameters such as the median x , y , and z location and the location of the 25th and 75th percentile x , y , and z position. Comparing the parameters generated from manual labels with parameters generated

from the MAP results, an average relative error of 4.98% was observed. Because the error rates for the individual features did not differ much, all experiments have been performed with a fixed reliability parameter for all conditions: $\gamma_1 = \gamma_2 = \dots = \gamma_F = \gamma$.

Following equations (7) to (10), using different values for γ results in different conditional SSMs. $\gamma = 0.0$ represents a SSM without conditions, $\gamma = 1.0$ represents the conventional conditional SSM and the relaxed conditional SSM is constructed in the domain $0.0 < \gamma < 1.0$. Within this range the optimal value of γ is searched with intervals of 0.1. $\gamma = 0.5$ showed best performance in generalization [12] for the relaxed SSM, when using the second 48 cases.

The results obtained by the proposed method were compared with the results obtained through ridge regression (see equations (11) and (12)). The ridge parameter ρ was optimized by searching for the maximum generalization within a range of $0.01 \leq \rho \leq 10000$, in which ρ was iteratively multiplied by a factor 10. The optimal ridge parameter was found for $\rho = 1000$, when using the second 48 cases.

Fig. 4a shows the degree of overlap between the estimated prior and the correct shape, for the 24 selected test cases. Compared to $\gamma = 0.0$ and $\gamma = 1.0$, a reliability parameter of $\gamma = 0.5$ generates the best results. The shape prior estimation for $\gamma = 0.5$

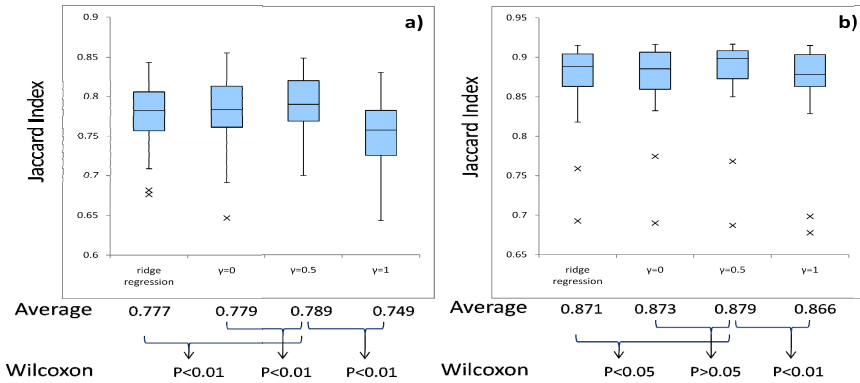


Fig. 4. Degree of overlap between shape prior and true shape (a) and between liver segmentation results and the true shape (b) for (from left to right) ridge regression and the proposed method with $\gamma = 0.0$, $\gamma = 0.5$ and $\gamma = 1.0$. Results for 24 difficult cases.

also outperforms shape prior estimation using ridge regression. Despite the small differences, the Wilcoxon signed rank test showed statistical significant differences. Fig. 5 shows an example result of a constructed shape prior. The relaxed conditional SSM ($\gamma = 0.5$) outperforms the other models. Notoriously difficult to segment areas, such as at the tip of the left lobe and at the bottom of the right lobe of the liver show better shape estimation results for the relaxed conditional SSM.

The Graph Cut based liver segmentation used the following energy function:

$$E(A) = \sum_{p \in P} \lambda \cdot R_p(A_p) + \sum_{(p,q) \in NB} \{B_{pq}(A_p, A_q) + S_{p,q}(A_p, A_q)\} \cdot \delta_{A_p \neq A_q}, \quad (13)$$

$$\delta_{A_p \neq A_q} = \begin{cases} 1 & (\text{if } A_p \neq A_q) \\ 0 & (\text{if } A_p = A_q) \end{cases}, \quad (14)$$

in which P is the set of voxels in CT images, $p \in P$ denotes the voxels, NB denotes the set of neighboring voxel pairs, $A = (A_1, \dots, A_p, \dots, A_{|P|})$ is the set of labels assigned to all voxels and λ is the weight factor to balance both energies. The other parameters are:

$$R_p(A_p) = \begin{cases} -\Pr(I_p | \text{"obj"})\Pr(\text{"obj"}) & (\text{if } A_p = \text{"obj"}) \\ -\Pr(I_p | \text{"bkg"})\Pr(\text{"bkg"}) & (\text{if } A_p = \text{"bkg"}) \end{cases}, \quad (15)$$

$$B_{pq}(A_p, A_q) = \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \frac{1}{\|p - q\|}, \quad (16)$$

$$S_{p,q}(A_p, A_q) = \sqrt{\frac{1}{2} \left(1 - \frac{\overline{pq} \cdot \nabla \Phi_p}{\|pq\| \|\nabla \Phi_p\|} \right)} \quad (17)$$

with I_p signifying the CT value of voxel p . Equation (15) calculates for every voxel the negative likelihood, equation (16) is the boundary term and equation (17) is the shape energy term. In this equation, Φ is the signed distance to the outline of the shape prior. By calculating the inner product of the vector from voxel p to neighboring voxel q with $\nabla \Phi_p$, the validity of the segmented shape is evaluated. Using the second set of 48 cases, optimized values were found at $\lambda = 1.5$ and $\sigma = 10.0$ (for ridge regression, for $\gamma = 0.0$ and for $\gamma = 1.0$), $\lambda = 1.0$ and $\sigma = 5.0$ (for $\gamma = 0.5$).

Fig. 4b shows the degree of overlap between the true shape and the result of Graph Cut segmentation with shape priors obtained from ridge regression and obtained from the proposed method with $\gamma = 0.0$, $\gamma = 0.5$ and $\gamma = 1.0$, when using the 24 difficult

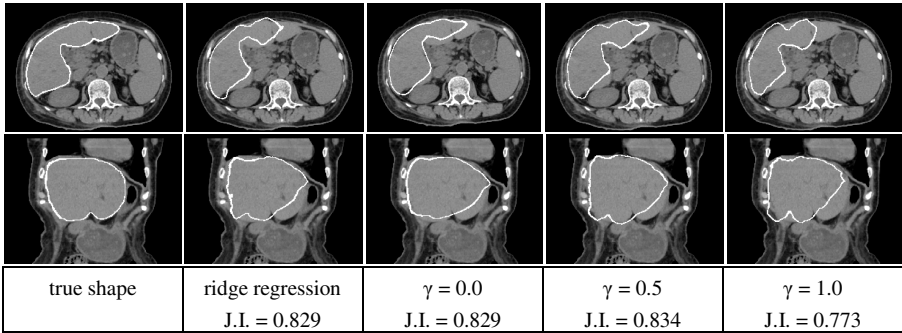


Fig. 5. Axial (top row) and coronal (bottom row) example results of generated shape priors for ridge regression and for the proposed method with different values of γ

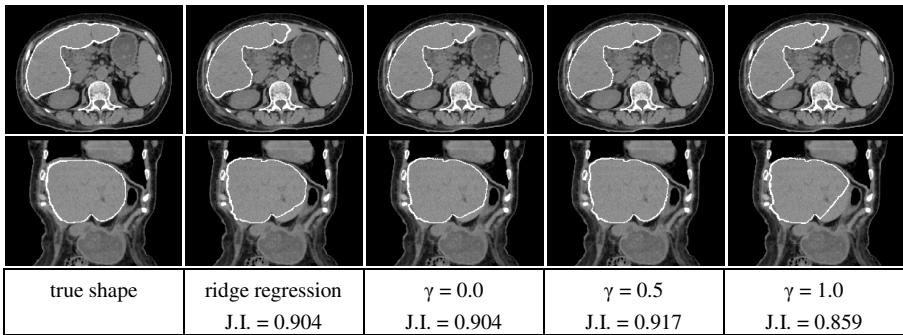


Fig. 6. Axial (top row) and coronal (bottom row) example segmentation results for Graph Cuts initialized with different shape priors

cases from the test data set. Wilcoxon signed rank test showed no statistical significant difference between the results for the standard SSM and the relaxed conditional SSM segmentation results. However, comparing the results obtained with the conventional and the relaxed SSM showed a statistical significant difference. Also, the proposed method ($\gamma = 0.5$) outperforms Graph Cut segmentation in which the shape prior was estimated using ridge regression, with a statistically significant difference in performance. Fig. 6 shows an example segmentation result for the four different models and, similar to Fig. 5, the relaxed conditional SSM generates the best results. The example image shows a liver with an extremely large left lobe. The axial view in Fig. 5 clearly shows the performance gain by the strongly improved segmentation of the tip of the left liver lobe. The coronal view shows that, contrary to the other models, the proposed relaxed conditional SSM is able to properly segment the bottom of left lobe, despite its far from average shape.

The 24 easy cases showed comparable results as the difficult cases, with differences however having weaker statistical significance. Average Jaccard Index values for ridge regression, $\gamma = 0.0$, $\gamma = 0.5$ and $\gamma = 1.0$ were 0.821, 0.820, 0.826 and 0.800 for shape estimation and 0.922, 0.929, 0.928 and 0.925 for subsequent Graph Cut segmentation. Average evaluation time for MAP based rough segmentation (Intel® Xeon® E5606 CPU), shape estimation (Nvidia® Tesla C2050 GPU) and Graph Cut segmentation (Intel® Xeon® E5606 CPU) were 20, 130 and 20 seconds respectively.

6 Discussion

This paper described the construction of a shape prior by a novel relaxed conditional SSM. The generated shapes were used as shape priors for Graph Cut segmentation of the liver in abdominal CT images. This way, the reliability of the condition is taken into account during the generation of the shape prior. A fixed reliability parameter γ was used for all conditions and was optimized using a data set of 48 training cases. Future work will focus on using different values of γ for individual features.

The goal of the research presented in this paper was to improve the accuracy of estimated shape priors. Compared to a conventional conditional SSM, compared to a SSM without condition and compared to a conditional SSM based on ridge

regression, the generated shape priors showed statistical significantly higher accuracy, for the 24 selected difficult to segment cases, thereby achieving the goal of this paper.

Because ridge regression is the only comparable method to bridge between the non-conditional SSM and the conventional conditional SSM, both using a conditional covariance matrix and a conditional average, the results of the proposed method have been compared with ridge regression based shape prior estimates. The proposed relaxed conditional SSM proved to estimate statistical significantly better shape priors.

Evaluating Fig. 4, the benefit of the proposed relaxed conditional SSM clearly lies in improved shape estimation and improved segmentation for difficult to segment images. Inspection of notoriously difficult to segment areas, such as shown in Figs. 5 and 6, corroborate the suggested benefit of the proposed method.

After Graph Cut segmentation, the results based on the relaxed conditional SSM still showed the highest Jaccard Index, also when compared with ridge regression based results. The proposed relaxed conditional SSM showed a higher average segmentation accuracy than all other models, with differences being statistically significant, except when compared to the non-conditional SSM. The improved accuracy in comparison with segmentation based on the conventional conditional SSM was found statistically significant. Therefore, it can be concluded that the relaxed conditional SSM outperforms the conventional conditional SSM, both in the estimation of the shape prior and in the subsequent segmentation.

The calculation of the conditional covariance matrix and the conditional average allows a seamless transition between the generic non-conditional SSM and the conventional conditional SSM. Contrary to [6], in which only a conditional covariance matrix is used, the proposed method calculates both a conditional covariance matrix and a conditional average. Following equations (7) to (10) a perfect interpolation between the non-conditional SSM and the conventional conditional SSM is achieved.

In ridge regression there is an over-accentuation of the weaker shape variations, which are more strongly influenced by the ridge parameter than the stronger shape variations. Such an imbalance does not arise when calculating the conditional covariance matrix and conditional average through equations (9) and (10). The improved performance in shape estimation and its influence on subsequent Graph Cut segmentation can be contributed to this seamless transition between the non-conditional SSM and the conventional conditional SSM.

Future work includes modifications to the algorithm, to enable the processing of contrast enhanced CT data as well. This will also allow a more thorough comparison with other methods, for example based on the SLIVER07 database [13].

Acknowledgements. Part of this research was performed under a Grant-in-aid for scientific research from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

1. Boykov, Y., Funka-Lea, G.: Graph Cuts and Efficient N-D Image Segmentation. *Int. J. Comput. Vis.* 70, 109–131 (2006)
2. Freedman, D., Zhang, T.: Interactive Graph Cut Based Segmentation with Shape Priors. In: *IEEE Computer Vision and Pattern Recognition*, pp. 755–762. IEEE Press, New York (2005)

3. Shimizu, A., Nakagomi, K., Narihira, T., Kobatake, H., Nawano, S., Shinozaki, K., Ishizu, K., Togashi, K.: Automated Segmentation of 3D CT Images Based on Statistical Atlas and Graph Cuts. In: Menze, B., Langs, G., Tu, Z., Criminisi, A. (eds.) MICCAI 2010. LNCS, vol. 6533, pp. 214–223. Springer, Heidelberg (2011)
4. de Bruijne, M., Lund, M.T., Tanko, L.B., Pettersen, P.C., Nielsen, M.: Quantitative Vertebral Morphometry Using Neighbor-Conditional Shape Models. *Med. Image Anal.* 11, 503–512 (2007)
5. Lord, F.M., Novick, M.R.: *Statistical Theories of Mental Test Scores*, pp. 146–147. Addison-Wesley Publishing Company Inc. (1968)
6. Baka, N., de Bruijne, M., Reiber, J.H.C., Niessen, W., Lelieveldt, B.P.F.: Confidence of Model Based Shape Reconstruction from Sparse Data. In: 7th IEEE International Symposium on Biomedical Imaging, pp. 1077–1080. IEEE Press, New York (2010)
7. Syrkina, E., Blanc, R., Szekely, G.: Propagating Uncertainties in Statistical Model Based Shape Prediction. In: Proc. of SPIE Medical Imaging, vol. 7962, p. 796240 (2011)
8. Hoerl, A., Kennard, R.: Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* 12, 55–67 (1970)
9. Leventon, M.E., Grimson, W.E.L., Faugeras, O.: Statistical shape influence in geodesic active contours. In: IEEE Computer Vision and Pattern Recognition, pp. 316–323 (2000)
10. Shimizu, A., Ohno, R., Ikegami, T., Kobatake, H., Nawano, S., Smutek, D.: Segmentation of Multiple Organs in Non-Contrast 3D Abdominal CT Images. *Int. J. Comput. Assist. Radiol. Surg.* 2, 135–142 (2007)
11. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: *Numerical Recipes*, pp. 509–514. Cambridge University Press (2007)
12. Styner, M.A., Rajamani, K.T., Nolte, L.-P., Zsemlye, G., Székely, G., Taylor, C.J., Davies, R.H.: Evaluation of 3D Correspondence Methods for Model Building. In: Taylor, C.J., Noble, J.A. (eds.) IPMI 2003. LNCS, vol. 2732, pp. 63–75. Springer, Heidelberg (2003)
13. Online Available at: <http://www.silver08.org>

A Landmark-Based Primal-Dual Approach for Discontinuity Preserving Registration

Silja Kiriyanthan, Ketut Fundana, Tahir Majeed, and Philippe C. Cattin

Medical Image Analysis Center, University of Basel,
CH-4031 Basel, Switzerland
silja.kiriyanthan@stud.unibas.ch,
{ketut.fundana,tahir.majeed,philippe.cattin}@unibas.ch

Abstract. Discontinuous motion is quite common in the medical field as for example in the case of breathing induced organ motion. Registration methods that are able to preserve discontinuities are therefore of special interest. To achieve this goal we developed in our previous work a framework that combines motion segmentation and registration. To avoid unreliable motion fields the incorporation of landmark correspondences can be a remedy. We therefore describe in this paper how we integrate the landmarks in our variational approach and how to solve the minimisation problem with a primal-dual algorithm. Qualitative and quantitative results are shown for real MR images of breathing induced liver motion.

Keywords: Motion, registration, liver.

1 Introduction

Nowadays, image registration is an indispensable tool for many medical applications and a great variety of well established methods have been proposed. Although image registration has been a topic of high interest in the last decades, methods that can handle discontinuous motion fields have only drawn little attention until recently. This despite its importance when for example registering abdominal organs that undergo breathing induced motion.

More attention to discontinuity preserving methods is being paid in the related research field of optical flow. Important theoretical contributions also originate from image segmentation and image denoising methods. Mumford and Shah for example proposed in their pioneering work [13] a functional for image segmentation that avoids spatial smoothing in certain locations of the image, thus preserving discontinuities. Vese and Chan [18] introduced a level set framework based approach to efficiently solve the Mumford and Shah minimisation problem for segmentation. Another influential approach based on the Total Variation (TV) norm, known to preserve discontinuities, was proposed by Rudin, Osher and Fatemi [16] for image denoising. The beneficial behaviour of the TV-norm was also exploited in image segmentation, image registration and optical flow methods, as for example in [5], [15] and [2].

A recent registration approach that tries to handle discontinuities in the displacement field of medical images has been proposed by Schmidt-Richberg *et al.* [17], which considers a direction-dependent regularisation method of the displacement field. This method relies on the calculation of the normals at the object boundaries and therefore a rather good manual segmentation has to be provided in advance.

As shown in the work of Amiaz *et al.* [1], where the optical flow method of Brox *et al.* [2] was embedded into the segmentation framework of Vese and Chan [18], the so achieved motion segmentation can influence the registration process positively. Instead of using the level set formulation, we used in our previous works [9,10] the segmentation framework of Chan *et al.* [5] that guarantees a globally optimal motion segmentation result for a fixed motion field. A preliminary short version of the work in [10] can be found in [11]. Following the work of Chambolle and Pock [4] we solved then the registration problem with a primal-dual approach.

To avoid unreliable motion fields, we incorporate in this paper the information of landmarks into our previous work [10]. There exist many image registration methods which integrate the information of landmark correspondences, as for example [7,8,12,14,3]. Here, we will make use of the recent work of Brox and Malik [3], which contains the idea on how to include descriptor matchings, respectively landmark correspondences, easily into a variational framework.

2 Method

In the following we want to describe the proposed method. First, we shortly recapitulate the registration and motion segmentation framework we used already before in [10] and in a second part we discuss the incorporation of the landmarks into the energy functional.

2.1 Registration and Motion Segmentation Framework

We define by $\Omega \subset \mathbb{R}^2$ the domain of the pixel positions $\mathbf{x} = (x_1, x_2)$ and by the functions $R : \Omega \rightarrow \mathbb{R}$ and $T : \Omega \rightarrow \mathbb{R}$ our reference and template image. The aim of image registration is to find a transformation $\Phi(\mathbf{x}) := \mathbf{x} + \mathbf{w}(\mathbf{x})$ such that the relation $T \circ \Phi \approx R$ holds and the displacement field $\mathbf{w} : \Omega \rightarrow \mathbb{R}^2$, where $\mathbf{w}(\mathbf{x}) := (u(\mathbf{x}), v(\mathbf{x}))$ with $u, v : \Omega \rightarrow \mathbb{R}$, will be the function we focus at. For convenience we will use the abbreviations \mathbf{w} , u and v for $\mathbf{w}(\mathbf{x})$, $u(\mathbf{x})$ and $v(\mathbf{x})$.

The proposed registration method integrates the displacement field estimation into the convex segmentation method of Chan *et al.* [5] to preserve the discontinuities in the displacement field. The energy functional for this variational approach is given by

$$E(\mathbf{w}^+, \mathbf{w}^-, \tilde{u}) = \int_{\Omega} D(\mathbf{w}^+) \tilde{u}(\mathbf{x}) d\mathbf{x} + \int_{\Omega} D(\mathbf{w}^-) (1 - \tilde{u}(\mathbf{x})) d\mathbf{x} + \nu \int_{\Omega} |\nabla \tilde{u}(\mathbf{x})| d\mathbf{x}. \quad (1)$$

Here, the function D represents the data term and is of the general form $D(\mathbf{w}) := f(\mathbf{w}) + \mu s(\mathbf{w})$, where f and s are the fidelity term and the smoothness term with $\mu \in \mathbb{R}^+$ being a weighting parameter. We choose a fidelity term f that incorporates the constraints for the grey value constancy and the gradient constancy with their corresponding weights $\gamma_1, \gamma_2 \in \mathbb{R}_0^+$. Therefore we define

$$f(\mathbf{w}) = f(u, v) := \gamma_1 |T(\mathbf{x} + \mathbf{w}) - R(\mathbf{x})| + \gamma_2 |\partial_{x_1} T(\mathbf{x} + \mathbf{w}) - \partial_{x_1} R(\mathbf{x})| \\ + \gamma_2 |\partial_{x_2} T(\mathbf{x} + \mathbf{w}) - \partial_{x_2} R(\mathbf{x})|. \quad (2)$$

The smoothness term s corresponds to the L^1 norm respectively the vectorial TV norm of \mathbf{w} and is given by

$$s(\mathbf{w}) = s(u, v) := \sqrt{|\nabla u|^2 + |\nabla v|^2} = |\nabla \mathbf{w}|. \quad (3)$$

Furthermore, to incorporate Chan *et al.*'s work [5], a binary function $\tilde{u} : \mathbb{R}^2 \rightarrow \{0, 1\}$, $\tilde{u}(\mathbf{x}) := \mathbf{1}_\Sigma(\mathbf{x})$, where $\Sigma \subseteq \Omega \subseteq \mathbb{R}^2$, with $\Sigma := \{\mathbf{x} \in \Omega \mid \tilde{u}(\mathbf{x}) = 1\}$, is used in (1) to differentiate the displacement field \mathbf{w} into \mathbf{w}^+ and \mathbf{w}^- . Finally, the last term in the above energy (1) is a regularisation defined by the TV norm and weighted by a parameter $\nu \in \mathbb{R}^+$.

The registration problem is solved by minimising energy E in (1) with respect to \mathbf{w}^+ , \mathbf{w}^- and \tilde{u} and we finally obtain the aimed displacement field by setting $\mathbf{w} := \mathbf{w}^+ \tilde{u} + \mathbf{w}^- (1 - \tilde{u})$.

As pointed out by Chan *et al.* in [5], (1) is strongly related to the Mumford-Shah functional [13] and one can show that a global minimiser of the set Σ can be found by minimising energy E in (1) with respect to \tilde{u} over a convex set.

2.2 Incorporation of the Landmarks

To exploit the information of the landmarks in the proposed variational registration method we make use of Brox and Malik's approach [3]. The landmark point correspondences, which are defined by the vectors $\mathbf{w}_{LM} = (u_{LM}, v_{LM})$, should act as a prior to the displacement field \mathbf{w} and we therefore seek to additionally minimise an energy of the form

$$E_{LM}(\mathbf{w}_{LM}) = \int_{\Omega} \mathbf{1}_S(\mathbf{x}) |\mathbf{w}(\mathbf{x}) - \mathbf{w}_{LM}(\mathbf{x})|_2^2 dx, \quad (4)$$

where S is the set of the landmark positions in the reference image R . More specifically, incorporating energy E_{LM} into our registration and motion segmentation framework defined by the energy functional E in (1) results in a modified fidelity term f in (2), namely

$$f(\mathbf{w}) := \gamma_1 |T(\mathbf{x} + \mathbf{w}) - R(\mathbf{x})| + \gamma_2 |\partial_{x_1} T(\mathbf{x} + \mathbf{w}) - \partial_{x_1} R(\mathbf{x})| \\ + \gamma_2 |\partial_{x_2} T(\mathbf{x} + \mathbf{w}) - \partial_{x_2} R(\mathbf{x})| + \frac{1}{2} \lambda \mathbf{1}_S(\mathbf{x}) |\mathbf{w}(\mathbf{x}) - \mathbf{w}_{LM}(\mathbf{x})|_2^2, \quad (5)$$

where λ is a weighting parameter. Instead of using a dense correspondence field [3], we use the landmarks extracted by the Affine Scale Invariant Feature Transform (A-SIFT) method [19], a recently developed algorithm that extends the well known SIFT method and allows landmark matching under affine deformations, hence usually finding a lot more matches than the SIFT method. The putative matches found are then filtered by fitting a homography to the matches using RANSAC yielding the set S of landmark positions.

3 Minimisation

3.1 Iterative Scheme

To facilitate the minimisation procedure we replace the fidelity term f in (5) by its partly linearised version

$$f(\mathbf{w}) = \gamma_1 |\rho_1(\mathbf{w})| + \gamma_2 |\rho_2^{(1)}(\mathbf{w})| + \gamma_2 |\rho_2^{(2)}(\mathbf{w})| + \frac{1}{2} \lambda \mathbf{1}_S(\mathbf{x}) \|\mathbf{w}(\mathbf{x}) - \mathbf{w}_{LM}(\mathbf{x})\|_2^2, \quad (6)$$

where

$$\rho_1(\mathbf{w}) := T(\mathbf{x} + \mathbf{w}_0) + \nabla T(\mathbf{x} + \mathbf{w}_0)^\top (\mathbf{w} - \mathbf{w}_0) - R(\mathbf{x}), \quad (7)$$

$$\rho_2^{(1)}(\mathbf{w}) := \partial_{x_1} T(\mathbf{x} + \mathbf{w}_0) + \begin{pmatrix} \partial_{x_1 x_1} T(\mathbf{x} + \mathbf{w}_0) \\ \partial_{x_2 x_1} T(\mathbf{x} + \mathbf{w}_0) \end{pmatrix}^\top (\mathbf{w} - \mathbf{w}_0) - \partial_{x_1} R(\mathbf{x}), \quad (8)$$

and

$$\rho_2^{(2)}(\mathbf{w}) := \partial_{x_2} T(\mathbf{x} + \mathbf{w}_0) + \begin{pmatrix} \partial_{x_1 x_2} T(\mathbf{x} + \mathbf{w}_0) \\ \partial_{x_2 x_2} T(\mathbf{x} + \mathbf{w}_0) \end{pmatrix}^\top (\mathbf{w} - \mathbf{w}_0) - \partial_{x_2} R(\mathbf{x}). \quad (9)$$

The minimisation of the energy functional E with respect to \mathbf{w}^+ , \mathbf{w}^- and \tilde{u} is then performed by the following iterative scheme:

1. For fixed \mathbf{w}^+ and \mathbf{w}^- , solve

$$\min_{\tilde{u} \in [0,1]} \left\{ \int_{\Omega} D(\mathbf{w}^+) \tilde{u}(\mathbf{x}) \, d\mathbf{x} + \int_{\Omega} D(\mathbf{w}^-) (1 - \tilde{u}(\mathbf{x})) \, d\mathbf{x} + \int_{\Omega} \nu |\nabla \tilde{u}(\mathbf{x})| \, d\mathbf{x} \right\}. \quad (10)$$

2. For fixed \tilde{u} , solve

$$\min_{\mathbf{w}^+} \left\{ \int_{\Omega} D(\mathbf{w}^+) \tilde{u}(\mathbf{x}) \, d\mathbf{x} \right\}. \quad (11)$$

3. For fixed \tilde{u} , solve

$$\min_{\mathbf{w}^-} \left\{ \int_{\Omega} D(\mathbf{w}^-) (1 - \tilde{u}(\mathbf{x})) \, d\mathbf{x} \right\}. \quad (12)$$

To solve the subproblems (10), (11) and (12) in a fast and efficient way, we follow a primal-dual approach as described by Chambolle and Pock in [4]. We therefore recapitulate in the next section the basic notations and formulations.

3.2 The Primal-Dual Approach of Chambolle and Pock

First, we define by X and Y two finite-dimensional real vector spaces. Their inner products are denoted by $\langle \cdot, \cdot \rangle_X$ respectively $\langle \cdot, \cdot \rangle_Y$ and their induced norms are given by $\|\cdot\|_X = \sqrt{\langle \cdot, \cdot \rangle_X}$ respectively $\|\cdot\|_Y = \sqrt{\langle \cdot, \cdot \rangle_Y}$. The general non-linear primal problem we have is of the form

$$\min_{x \in X} F(Kx) + G(x), \tag{13}$$

where $F : Y \rightarrow [0, +\infty)$ and $G : X \rightarrow [0, +\infty)$ are proper, convex and lower semi-continuous and the map $K : X \rightarrow Y$ is a continuous linear operator. The corresponding primal-dual formulation is the saddle-point problem

$$\min_{x \in X} \max_{y \in Y} \langle Kx, y \rangle_Y + G(x) - F^*(y), \tag{14}$$

with $F^* : Y \rightarrow \mathbb{R} \cup \{+\infty\}$ being the convex conjugate of F . We assume that the problems above have at least one solution $(\hat{x}, \hat{y}) \in X \times Y$ and therefore it holds $K\hat{x} \in \partial F^*(\hat{y})$ and $-(K^*\hat{y}) \in \partial G(\hat{x})$, where $\partial F^*(\hat{y})$ and $\partial G(\hat{x})$ are the subdifferentials of the convex functions F^* at \hat{y} and G at \hat{x} . Furthermore we assume that F and G are “simple”, *i.e.* that the resolvent operators $(I + \sigma \partial F^*)^{-1}$ and $(I + \tau \partial G)^{-1}$ are easy to compute. For a convex function f the resolvent of the operator $\tau \partial f$ at \tilde{x} can be calculated in our case by

$$x = (I + \tau \partial f)^{-1}(\tilde{x}) = \arg \min_x \left\{ \frac{\|x - \tilde{x}\|^2}{2\tau} + f(x) \right\}. \tag{15}$$

We will only make use of Algorithm 1 in [4] with the extrapolation parameter $\theta = 1$. The usage of the other proposed algorithms is left for the moment for later research.

To apply Algorithm 1 in [4] to the minimisation problems (10), (11) and (12), we first need to rewrite them in their discretised version, then identify the functions F and G and finally derive the resolvent operators $(I + \sigma \partial F^*)^{-1}$ and $(I + \tau \partial G)^{-1}$.

For the discrete setting we therefore define by $\mathbf{x}_{i,j} = (x_{1,i,j}, x_{2,i,j}) = (ih, jh)$, $i = 1, \dots, M$, $j = 1, \dots, N$, the pixel positions in the image domain with h being the spatial step size. For the calculations of the finite differences, the discrete divergence operator, the discretised inner products and further details we refer the reader to [4] and the references therein.

In the following sections we will discuss the resolvent operators for the three given minimisation problems (10), (11) and (12). The formulation of the resolvent operators does not change much with respect to our previous work in [10]. We nevertheless recapitulate them in the following sections and point out the certain changes which appear.

3.3 Resolvent Operators for Problem (10)

Let us consider the continuous problem (10). After its discretisation and some calculations we get

$$p = (I + \sigma \partial F^*)^{-1}(\tilde{p}) \implies p_{i,j} = \nu \frac{\tilde{p}_{i,j}}{\max\{\nu, |\tilde{p}_{i,j}|\}}, \quad (16)$$

as a solution of the resolvent operator with respect to F^* . This is the same result we got in (10) and the calculations are performed in the same way as there. The resolvent operator with respect to G is also derived similar as in (10) and is given by

$$\tilde{u} = (I + \tau \partial G)^{-1}(\hat{u}) \implies \tilde{u}_{i,j} = \min\{\max\{\hat{u}_{i,j} - \tau (D(\mathbf{w}_{i,j}^+) - D(\mathbf{w}_{i,j}^-)), 0\}, 1\}. \quad (17)$$

Note that the data term D , or more specifically the fidelity term f that appears in D , is not the same as in our previous work. This time it additionally incorporates the information of the landmarks.

3.4 Resolvent Operators for Problem (11)

Now we consider the continuous problem (11). The numerical calculations in the implementation get facilitated by having a smooth extension of \mathbf{w}^+ to the domain $\Omega \setminus \Sigma$. We therefore consider instead the problem

$$\min_{\mathbf{w}^+} \left\{ \int_{\Omega} f(\mathbf{w}^+) \tilde{u}(\mathbf{x}) + \mu s(\mathbf{w}^+) d\mathbf{x} \right\}. \quad (18)$$

Comparing (11) to (18) the only difference is, that the factor \tilde{u} is not applied to the smoothness term s anymore.

From the resolvent operator with respect to F^* we obtain this time

$$q = (I + \sigma \partial F^*)^{-1}(\tilde{q}) \implies q_{i,j} = \mu \frac{\tilde{q}_{i,j}}{\max\{\mu, |\tilde{q}_{i,j}|\}}. \quad (19)$$

This is again the same result as in (10) but more changes appear instead in the solution for resolvent operator with respect to G . This time the function G incorporates the information of the landmarks and is given by

$$G(\mathbf{w}^+) = \sum_{i,j} \left(\gamma_1 |\rho_1(\mathbf{w}_{i,j}^+)| + \gamma_2 |\rho_2^{(1)}(\mathbf{w}_{i,j}^+)| + \gamma_2 |\rho_2^{(2)}(\mathbf{w}_{i,j}^+)| + \frac{1}{2} \lambda \mathbf{1}_S(\mathbf{x}_{i,j}) ((u_{i,j}^+ - u_{LM\ i,j})^2 + (v_{i,j}^+ - v_{LM\ i,j})^2) \right) \tilde{u}_{i,j}. \quad (20)$$

The derivation of the resolvent operator with respect to G is again not that straightforward and more effort has to be put in to find a suitable solution. This

can be done similar as in [10] and therefore, having a closer look at the definition of G (20) and equation (15), we see that we have to solve

$$\begin{aligned} \mathbf{w}^+ &= (I + \tau \partial G)^{-1}(\mathbf{w}_0^+) = \arg \min_{\mathbf{w}^+ \in X} \left\{ \frac{\|\mathbf{w}^+ - \mathbf{w}_0^+\|_X^2}{2\tau} + G(\mathbf{w}^+) \right\} \\ &= \arg \min_{\mathbf{w}^+ = (u^+, v^+) \in X} \left\{ \frac{1}{2\tau} \sum_{i,j} \left((u_{i,j}^+ - u_{0,i,j}^+)^2 + (v_{i,j}^+ - v_{0,i,j}^+)^2 \right) \right. \\ &\quad + \sum_{i,j} \left(\gamma_1 |\rho_1(\mathbf{w}_{i,j}^+)| + \gamma_2 |\rho_2^{(1)}(\mathbf{w}_{i,j}^+)| + \gamma_2 |\rho_2^{(2)}(\mathbf{w}_{i,j}^+)| \right. \\ &\quad \left. \left. + \frac{1}{2} \lambda \mathbf{1}_S(\mathbf{x}_{i,j}) \left((u_{i,j}^+ - u_{LM,i,j})^2 + (v_{i,j}^+ - v_{LM,i,j})^2 \right) \right) \tilde{u}_{i,j} \right\}. \end{aligned} \tag{21}$$

Similar to the work in [10] for $\tilde{u}_{i,j} = 0$ we can conclude from (21) that $\mathbf{w}_{i,j}^+ = \mathbf{w}_{0,i,j}^+$. On the other hand, for $\tilde{u}_{i,j} = 1$ we have again to distinguish the cases

$$\rho_1(\mathbf{w}_{i,j}^+) \begin{pmatrix} \geq \\ \leq \end{pmatrix} 0, \quad \rho_2^{(1)}(\mathbf{w}_{i,j}^+) \begin{pmatrix} \geq \\ \leq \end{pmatrix} 0, \quad \rho_2^{(2)}(\mathbf{w}_{i,j}^+) \begin{pmatrix} \geq \\ \leq \end{pmatrix} 0, \tag{22}$$

which turn out to be 27 in total. The additional term of the landmarks will cause slight changes in the resolvent operator with respect to G that we used to have before in [10]. Nevertheless, the idea of the derivation of the explicit solutions for $\mathbf{w}_{i,j}^+$ and the reformulations of the conditions in (22) remains the same.

3.5 Resolvent Operators for Problem (12)

This section is very similar to Section 3.4 and we mainly have to just replace \mathbf{w}^+ by \mathbf{w}^- and \tilde{u} by $1 - \tilde{u}$. To achieve a smooth extension of \mathbf{w}^- to the domain Σ , we consider now instead of problem (12) the following problem

$$\min_{\mathbf{w}^-} \left\{ \int_{\Omega} f(\mathbf{w}^-) (1 - \tilde{u}(\mathbf{x})) + \mu s(\mathbf{w}^-) dx \right\}. \tag{23}$$

The resolvent operator with respect to F^* is identical to (19) and compared to the section before only very slight changes of the resolvent operator with respect to G have to be done.

4 Results

The minimisation problems for the displacement fields \mathbf{w}^+ and \mathbf{w}^- remain non-convex, although we have a convex minimisation problem with respect to the motion segmentation function \tilde{u} . To facilitate the minimisation procedure of this overall non-convex problem, the fidelity term f was linearised in (6). Nevertheless we should remember that we have to update \mathbf{w}_0^+ and \mathbf{w}_0^- regularly. To avoid the risk of getting stuck in a local minimum during the optimisation we therefore

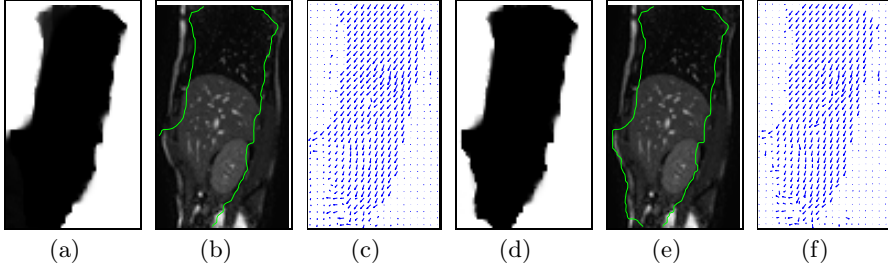


Fig. 1. Qualitative example for a pair of liver images with a discontinuous displacement field. Motion segmentation \hat{u} without landmarks (a) and with (d), reference image R overlaid with the motion segmentation without (b) and with landmarks (e) and the displacement field w without (c) and with landmarks (f).

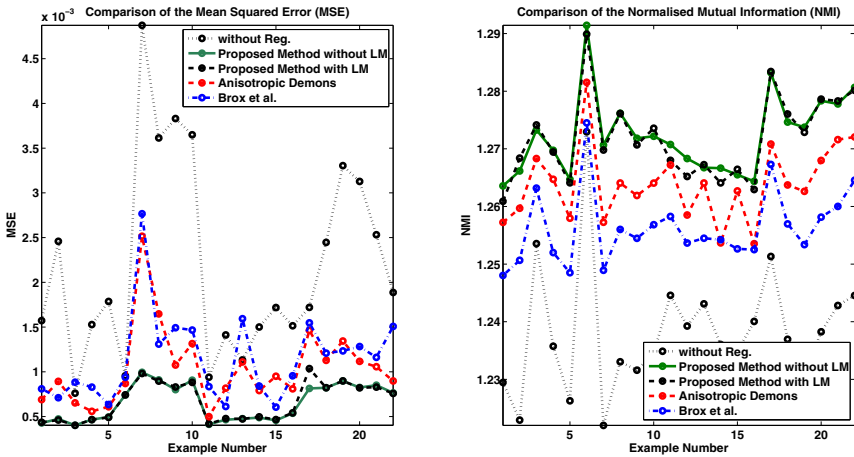


Fig. 2. Quantitative evaluation for 22 pairs of liver images with a discontinuous displacement field. Comparison of the MSE (left) and NMI (right).

apply a coarse-to-fine strategy in the same manner as in our previous work [10]. The final displacement field is then achieved by setting $w(x) = w^+(x)$ if $x \in \Sigma$ and $w(x) = w^-(x)$ if $x \in \Omega \setminus \Sigma$. To calculate the images $T(x + w^\pm)$ during the iterations and to obtain the final registered image $T(x + w)$ bicubic interpolation is used.

For the experiments we used real MR images of the abdomen, which were taken during the breathing cycle and show the sliding motion of the liver. A qualitative example is shown in Fig. 1. The inclusion of the landmarks lead to a more reliable motion segmentation result, since the in fact static part of the background is not assigned to the area with bigger motion anymore.

In Fig. 2 a quantitative evaluation is shown for 22 different liver image pairs. In average, the A-SIFT method [19] delivered around 1173 matches for these image pairs and after applying RANSAC around 1001 were left. Since we are interested in the discontinuities of the displacement field, we compare the proposed method to the methods that preserve discontinuities in the displacement field, in this

case, the demon algorithm with anisotropic diffusion filtering [6], the registration algorithm of Brox *et al.* [2] and our previous work [10]. The parameters for all the methods were chosen by optimising them with respect to the 22 image pairs. For our methods they were set to $\gamma_1 = 4$, $\gamma_2 = 1$, $\mu = 0.2$, $\nu = 0.1$ and for the weighting of the landmark term we used $\lambda = 0.3$. For the demon algorithm with anisotropic diffusion filtering we could use the suggested parameters and for Brox *et al.*'s method we used $\gamma = 5$, $\alpha = 80$ and $\sigma = 0.9$. Both of our methods showed an improvement compared to the demon algorithm with anisotropic diffusion filtering and the registration algorithm of Brox *et al.*

5 Conclusion

In this paper we presented a primal-dual method for discontinuity preserving non-rigid registration, that makes use of the segmentation framework of Chan *et al.* [5] and includes the information of landmarks. The so gained motion segmentation influences the motion estimation positively by sharpening the discontinuities in the displacement field. The minimisation of the energy functional is implemented in a coarse-to-fine strategy and exploits the rapidity of the primal-dual algorithm studied in [4]. The experimental results demonstrated desirable performance of the proposed method in comparison with those of the demon algorithm with anisotropic diffusion filtering [6] and the registration algorithm of Brox *et al.* [2].

Acknowledgements. This work has been funded by Dr. Hansjörg Wyss.

References

1. Amiaz, T., Kiryati, N.: Piecewise-Smooth Dense Optical Flow via Level Sets. *International Journal of Computer Vision* 68(2), 111–124 (2006)
2. Brox, T., Bruhn, A., Papenbergh, N., Weickert, J.: High Accuracy Optical Flow Estimation Based on a Theory for Warping. In: Pajdla, T., Matas, J.(G.) (eds.) *ECCV 2004*. LNCS, vol. 3024, pp. 25–36. Springer, Heidelberg (2004)
3. Brox, T., Malik, J.: Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33(3), 500–513 (2011)
4. Chambolle, A., Pock, T.: A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *Journal of Mathematical Imaging and Vision* 40(1), 120–145 (2011)
5. Chan, T.F., Esedoglu, S., Nikolova, M.: Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models. *Siam Journal on Applied Mathematics* 66(5), 1632–1648 (2006)
6. Demirovic, D., Serifovic, A., Cattin, P.C.: An anisotropic diffusion regularized demons for improved registration of sliding organs. In: *18th International Electrotechnical and Computer Science Conference, ERK* (2009)
7. Haber, E., Heldmann, S., Modersitzki, J.: A Scale-Space Approach to Landmark Constrained Image Registration. *Scale Space and Variational Methods in Computer Vision* 5567, 612–623 (2009)

8. Johnson, H., Christensen, G.: Consistent Landmark and Intensity-Based Image Registration. *IEEE Transactions on Medical Imaging* 21(5), 450–461 (2002)
9. Kiriyanthan, S., Fundana, K., Cattin, P.C.: Discontinuity Preserving Registration of Abdominal MR Images with Apparent Sliding Organ Motion. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) *Abdominal Imaging 2011*. LNCS, vol. 7029, pp. 231–239. Springer, Heidelberg (2012)
10. Kiriyanthan, S., Fundana, K., Majeed, T., Cattin, P.C.: Discontinuity Preserving Registration through Motion Segmentation: A Primal-Dual Approach (submitted)
11. Kiriyanthan, S., Fundana, K., Majeed, T., Cattin, P.C.: A Primal-Dual Approach for Discontinuity Preserving Registration. In: 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 350–353 (May 2012)
12. Kybic, J., Unser, M.: Fast Parametric Elastic Image Registration. *IEEE Transactions on Image Processing* 12(11), 1427–1442 (2003)
13. Mumford, D., Shah, J.: Optimal Approximations by Piecewise Smooth Functions and Associated Variational-Problems. *Communications on Pure and Applied Mathematics* 42(5), 577–685 (1989)
14. Paquin, D., Levy, D., Xing, L.: Hybrid Multiscale Landmark and Deformable Image Registration. *Mathematical Biosciences and Engineering* 4(4), 711–737 (2007)
15. Pock, T., Urschler, M., Zach, C., Beichel, R.R., Bischof, H.: A Duality Based Algorithm for $TV-L^1$ -Optical-Flow Image Registration. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part II*. LNCS, vol. 4792, pp. 511–518. Springer, Heidelberg (2007)
16. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear Total Variation Based Noise Removal Algorithms. *Physica D* 60(1-4), 259–268 (1992)
17. Schmidt-Richberg, A., Ehrhardt, J., Werner, R., Handels, H.: Slipping Objects in Image Registration: Improved Motion Field Estimation with Direction-Dependent Regularization. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part I*. LNCS, vol. 5761, pp. 755–762. Springer, Heidelberg (2009)
18. Vese, L.A., Chan, T.F.: A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model. *J. of Computer Vision* 50(3), 271–293 (2002)
19. Yu, G., Morel, J.M.: ASIFT: An Algorithm for Fully Affine Invariant Comparison. *Image Processing On Line* (2011), doi: <http://dx.doi.org/10.5201/ipol.2011.my-asift>

Predicting Liver Motion Using Exemplar Models

Golnoosh Samei, Christine Tanner, and Gabor Székely

Computer Vision Laboratory, Sternwartstrasse 7,
ETH Zentrum, 8092 Zürich, Switzerland
samei@vision.ee.ethz.ch

Abstract. The effective modeling and predicting of respiratory motion in abdominal organs is crucial to the task of tumor treatments. Current approaches in statistical respiration modeling either build a subject-specific model which is only suitable for that single subject or create a population model by assuming a coherent population with a relatively simple distribution and, therefore, fail to account for variations of the breathing pattern among different subjects. To bridge this gap, we propose a more flexible method based on exemplar models, which is able to cope with heterogeneous population data and can be better adapted to a previously unseen subject. We have showed that, in contrary to principal component analysis based models, our method is capable of effectively utilizing complementary information provided by increasing number of examples taken from a population. In addition to being more robust against outliers, the proposed method also achieved lower mean errors in leave-one-out experiments.

Keywords: Respiration, liver motion, statistical population model, motion modeling, 4D MRI, principal component analysis, exemplar models.

1 Introduction

To take advantage of novel therapy devices with improved targeting precision (e.g. beam-steering, pencil-beams, phased arrays, refined catheter-steering, multileaf collimators), it is crucial to have accurate information about the position of the organ during therapy. Fast and comprehensive image acquisition as well as fast image processing would be required for determining and quantifying the motion of all points in the organ in real time. As this is generally not possible, prior knowledge about the expected motion is required to complement the partial observations (so called surrogates) acquired during therapy.

There is an extensive body of work dedicated to modeling 4D motion in abdominal organs, and the interested reader is referred to [1] for a review on this topic. These works follow two general approaches. Either a so called subject-specific model is built from the data acquired through a planning session [2–4], or data from different subjects are gathered to build a population model [5–10]. The former has the advantage that the model captures the specific motion patterns of an individual subject. However, any inter-session changes in the organ motion are difficult to account for in this approach. Moreover, building subject-specific

models could be time consuming and expensive. Population models are designed to overcome these shortcomings. In most cases these are based on principal component analysis (PCA) [5, 8, 10, 11]. An exception is the work of Ehrhardt et al. [9], where a mean motion model was built for lung cancer patients by mapping all subject-specific motion models to an average image space and calculating the Log-Euclidean mean of all the mapped transformations. Von Siebenthal et al. [5] observed that, in addition to the repetitive respiratory motion of the liver, there exists a gradual motion and deformation of the organ which is noticeable in imaging sessions lasting more than 20 minutes [5]. They built a population model for this so called drift in the observed livers using PCA. Most recently, Arnold et al. [11] used this drift model in combination with a subject-specific recording of liver positions over time for spatial as well as temporal prediction.

However, all the aforementioned population models assume a coherent population (e.g. one Gaussian distribution modeled by PCA), which is unlikely to hold as there are variations in the motion patterns of different subjects. He et al. [6] used kernel PCA with least-square support vector machines to model the non-linear statistics of respiratory motion. However, they did not specifically address the issue of population sub-classes. Even though their model is more powerful than PCA, the tuning of the parameters of the kernel function as well as the support vector algorithm is a time consuming and cumbersome process and might have to be repeated for each new dataset.

To address these issues, we propose the use of exemplar models as a non-parametric method for adapting the population model to an individual subject during therapy based on the individual breathing patterns observed via the surrogate. To that end, we have created submodels by PCA analysis for individual subjects. The final model is a linear mixture of these exemplar models. Even though the use of most similar samples has been previously proposed in other fields such as atlas-based segmentations [12], to the best of our knowledge this is the first attempt to build an individualized breathing model from observations over a population by taking advantage of similarities in the respiratory motion of different subjects.

2 Materials

Our experiments are based on free-breathing 4DMRI [5] data obtained from 12 healthy volunteers, 6 female and 6 male, with an average age of 31. The 4DMRI sequences consist of 25-30 slices with an in-plane resolution of $1.8 \times 1.8 \text{ mm}^2$, a slice thickness of 4 mm and a temporal resolution of 290-410 ms taken in the course of 40-75 minutes.

An intensity-based non-rigid registration has been employed [5] to quantify the motion of the liver observed during 4DMRI acquisition. In order to establish inter-subject correspondences, a number of anatomically and biomechanically corresponding points were chosen manually for each subject. By performing a cubic interpolation between these landmarks, the positions of 290 corresponding points in the liver were obtained [5].

In this study we used the position of three of these points as our surrogates and assumed that they can be tracked during therapy. These surrogates include a point on the diaphragm, the entrance point of the portal vein into the liver, and a point in the center of the liver defined by vessel features. We also considered the case where the liver motion of a subject can be observed for 3-5 minutes immediately prior to therapy for creating a subject-specific respiratory model. We refer to these data, which typically consist of 27-84 respiratory cycles, as the pre-therapy data.

Motion Data. The liver position is described by N points, where the position of the i th point at time step t is represented by $\mathbf{p}_t^i = [p_{SI_t}^i, p_{AP_t}^i, p_{LR_t}^i]$, with SI, AP, and LR denoting the superior-inferior, anterior-posterior, and left-right axis. The position of the liver at time step t is represented by a $3N$ dimensional vector \mathbf{p}_t built from concatenating the \mathbf{p}_t^i vectors. The motion of the liver at time t is defined by $\Delta\mathbf{p}_t = \mathbf{p}_t - \mathbf{p}_{ref}$ with \mathbf{p}_{ref} being the reference exhale vector constructed using the available exhale image(s) during the pre-therapy stage. The exhale state t_{exh} in each cycle is defined as the local maxima of SI component of the mean motion trajectory of the observed points (i.e. surrogates). Motivated by [5], we decompose the generic motion vector $\Delta\mathbf{p}_t$ into respiratory $\Delta\mathbf{p}_{resp,t}$ and drift $\Delta\mathbf{p}_{drift,t}$ components:

$$\Delta\mathbf{p}_t = \Delta\mathbf{p}_{resp,t} + \Delta\mathbf{p}_{drift,t} = (\mathbf{p}_t - \mathbf{p}_{t_{exh}}) + (\mathbf{p}_{t_{exh}} - \mathbf{p}_{ref}) , \quad (1)$$

where $\mathbf{p}_{t_{exh}}$ is the position of the points at the most recent exhale state. The same decomposition is done for the surrogate signal during therapy, with each component of the surrogate being used to predict the respective component of the motion vector.

3 Methods

3.1 Motion Model Concept

PCA has been successfully employed in the literature for modeling the subject-specific respiratory motion of abdominal organs [3, 4] and the underlying reason of its effectiveness for this purpose has theoretically been explained [13]. Yet, PCA’s suitability for modeling a population of subjects is not as well justified. Indeed, variations in anatomy and breathing regimes between individual subjects are likely to lead to different motion patterns that are not well approximated by a single Gaussian distribution.

In this work we propose to distinguish between different patterns of breathing motion within a population. On the one hand, effective clustering of the motion vectors for such high dimensional data is not easily attained. On the other hand, it is reasonable to assume that motion vectors originating from an individual subject acquired during a single session demonstrate similar breathing patterns. Therefore, instead of finding different patterns of breathing motion by means of clustering, we turn to a classification method called instance-based

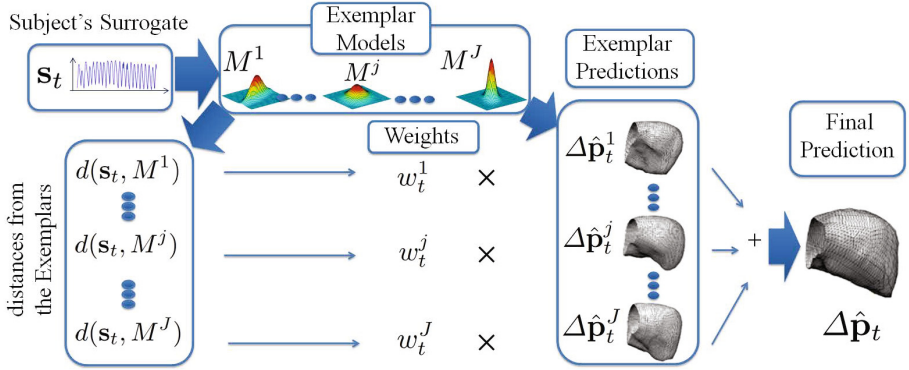


Fig. 1. This schematic illustrates the different steps of our proposed exemplar method for motion prediction during therapy

learning [14], which has been effectively exploited in computer vision and machine learning [15, 16].

In particular, our method is based on the distance-weighted k -nearest neighbor algorithm [17]. We have assumed that different examples of breathing patterns exist in the population and each of them is represented by at least one subject in the dataset. Hence, we built our exemplar models by fitting a PCA model to the motion vectors of each individual subject. During therapy, the prediction is realized as a weighted combination of the predictions of all the exemplar models, where the weights are based on the similarity of the surrogate and the corresponding model. The different steps of our proposed method are illustrated in Fig. 1.

3.2 Subject-Specific PCA Modeling

Assuming that the $3N$ dimensional motion vectors $\Delta \mathbf{p}_t, t = 1..T$, with T being the number of time steps, belong to a $3N$ dimensional Gaussian distribution $\Delta \mathbf{p}_t \sim \mathcal{N}(\mu, \Sigma)$, our task is to find the most probable vector $\Delta \hat{\mathbf{p}}_t$, given a subset of its elements, namely the surrogate vector \mathbf{s}_t . To create the motion model, the observed vectors $\Delta \mathbf{p}_t$ are decomposed into the surrogate \mathbf{s}_t and the motion vector of the rest of the points which we wish to predict (\mathbf{r}_t), as $\Delta \mathbf{p}_t^T = [\mathbf{s}_t^T, \mathbf{r}_t^T]$. The same decomposition into corresponding components is done for their mean $\mu^T = [\mu_s^T, \mu_r^T]$, and their covariance matrix $\Sigma = \begin{bmatrix} \Sigma_{ss} & \Sigma_{sr} \\ \Sigma_{rs} & \Sigma_{rr} \end{bmatrix}$. It can be shown [18], that if the distribution of the $\Delta \mathbf{p}_t$ vectors is Gaussian, then the conditional distribution ($\Delta \mathbf{p}_t | \mathbf{s}_t$) is also a Gaussian distribution of the form

$$(\Delta \mathbf{p}_t | \mathbf{s}_t) \sim \mathcal{N}\left(\mu + \begin{bmatrix} \Sigma_{ss} \\ \Sigma_{rs} \end{bmatrix} \Sigma_{ss}^{-1} (\mathbf{s}_t - \mu_s), \Sigma - \begin{bmatrix} \Sigma_{ss} \\ \Sigma_{rs} \end{bmatrix} \Sigma_{ss}^{-1} \begin{bmatrix} \Sigma_{ss} \\ \Sigma_{rs} \end{bmatrix}\right). \quad (2)$$

Therefore, the most probable vector $\Delta\hat{\mathbf{p}}_t$ given \mathbf{s}_t , is the mean of the above conditional distribution.

To find the principal components of the $\Delta\mathbf{p}_t$ vectors, we perform Eigenvalue decomposition on their covariance matrix, $\Sigma \stackrel{Eig}{=} \mathbf{E}\mathbf{\Lambda}\mathbf{E}^T$, where \mathbf{E} is composed of the eigenvectors of Σ , and $\mathbf{\Lambda}$ has the corresponding eigenvalues on its diagonal. By breaking up $\mathbf{E}^T = [\mathbf{E}_s^T, \mathbf{E}_r^T]$ into two submatrices, each consisting of the rows corresponding to \mathbf{s} and \mathbf{r} , we find the most probable vector $\Delta\hat{\mathbf{p}}_t$ given \mathbf{s}_t in terms of eigenvectors as:

$$\Delta\hat{\mathbf{p}}_t = \mu + \mathbf{E}\mathbf{\Lambda}\mathbf{E}_s^T(\mathbf{E}_s\mathbf{\Lambda}\mathbf{E}_s^T)^{-1}(\mathbf{s}_t - \mu_s) . \quad (3)$$

3.3 Population PCA Modeling

Given a dataset with J subjects, where each subject $j = 1..J$, has T_j observations of the same N points, a PCA population model is built similar to section 3.2, by using the observations from all included subjects j . Index t then denotes the observation index and ranges from 1 to $\sum_j T_j$. The rest of the algorithm is the same as the subject-specific algorithm.

3.4 Exemplars Modeling

To create a population model using exemplars, we build a PCA model M^j from each subject j as described in section 3.2. To predict $\Delta\mathbf{p}_t$, for a new subject, given \mathbf{s}_t at time step t , we obtain the motion vector predictions $\Delta\hat{\mathbf{p}}_t^j$ of all M^j models through equation 3. To combine these predictions in a distance-weighted k-nearest-neighbor approach [17], we have to estimate the dissimilarity of each of the exemplar models to the current observation. We base this dissimilarity on the squared Mahalanobis distance between the surrogate \mathbf{s}_t and the corresponding components of an individual model M^j :

$$d(\mathbf{s}_t, M^j) = (\mathbf{s}_t - \mu_s^j)^T \Sigma_{ss}^j (\mathbf{s}_t - \mu_s^j) . \quad (4)$$

Consequently, the exemplar weight, w_t^j , which provides the contribution of the prediction of model M^j at time t to the final prediction, is computed from the inverse of the above distance:

$$w_t^j = \frac{1/(d(\mathbf{s}_t, M^j) + \eta)}{\sum_{k=1}^J 1/(d(\mathbf{s}_t, M^k) + \eta)} , \quad (5)$$

where η is a small positive number introduced to avoid numerical instability. Finally, the motion vector, $\Delta\hat{\mathbf{p}}_t$, is formed from a weighted sum of the $\Delta\hat{\mathbf{p}}_t^j$ vectors:

$$\Delta\hat{\mathbf{p}}_t = \sum_{j=1}^J w_t^j \Delta\hat{\mathbf{p}}_t^j . \quad (6)$$

4 Experiments and Results

4.1 Evaluation Method

The prediction error of the model at each point \mathbf{p}_t^i was defined by the Euclidean distance between the predicted motion ($\Delta\hat{\mathbf{p}}_t^i$) and the motion obtained from the non-rigid registration result ($\Delta\mathbf{p}_t^i$), ie. $E_t^i = \|\Delta\hat{\mathbf{p}}_t^i - \Delta\mathbf{p}_t^i\|$. For each subject j , the error distribution was summarized by its mean mE_j , standard deviation sE_j and maximum xE_j , e.g.

$$mE_j = \frac{1}{T_j N} \sum_{t=1}^{T_j} \sum_{i=1}^N E_t^i . \quad (7)$$

The number of points in the liver of each subject was 290 and the number of time steps T_j varied between 3300 and 6700. The prediction accuracy of the population models were evaluated based on leave-one-out experiments. We used 3 of the points as our surrogates (see section 2) and predicted the positions of the remaining $N = 287$ points in the livers. No temporal prediction was performed.

Statistical significance at the $\alpha=0.001$ level of the difference in the mean of mE_j over all subjects of two methods was tested by bootstrapping [19]. For this, prediction results were resampled with replacement 5000 times and the difference in the mean error was calculated. The probability of equal results was based on the fraction of differences which were zero or negative.

4.2 Experiments

In this study, we performed two sets of experiments to evaluate two aspects of our proposed method, namely its prediction accuracy and its learning capability. In the following we will explain each set of experiments in more details.

Prediction Accuracy Experiments. To fully investigate the contribution of individual steps of our framework to the prediction accuracy, we conducted 7 experiments, with different combinations of the type of motion vectors {drift, respiration and generic = drift + respiration} and the type of motion model {exemplar models, subject-specific PCA and population PCA}. PCA and exemplar models of drift and respiratory motion were built by replacing $\Delta\mathbf{p}_t$ with $\Delta\mathbf{p}_{drift,t}$ and $\Delta\mathbf{p}_{resp,t}$ vectors, respectively, in the calculations of sections 3.2 and 3.4. Subject-specific PCA models were based on the pre-therapy data from the same subject. Since no substantial drift could be observed during the pre-therapy stage, no subject-specific drift motion models were created.

Learning Capacity Experiments. To test the ability of our proposed method to fully utilize the information provided by an increasing number of samples used for model building, we performed the following experiment. After excluding the subject to be tested, we randomly picked $n \in \{1..11\}$ subjects from the

population, and built a PCA and an exemplar drift model. Next, we used these models to predict the drift of the left out subject. The same procedure was carried out for all the subjects and the mean error over all of them, ε_n , was computed. We repeated this experiment 7 times for each number n and computed the mean and the standard deviation of ε_n over all 7 repetitions.

4.3 Results

Table 1 lists the statistics of the motion amplitude for the individual subjects. The results of the prediction accuracy experiments, using the generic motion models and the experiments using the separated motion models are summarized in Tables 2 and 3, respectively. These results suggest that,

- Exemplar models outperform the PCA models in all of the experiments (G3 vs G2, S2 vs S1, and S4 vs S3).
- Separation of drift and respiratory motion lowers the error only when the respiratory component is modeled by a subject-specific model (S3 and S4 have lower errors than G2 and G3, but S1 and S2 have higher errors than G2 and G3 respectively).
- The best results are achieved by modeling the drift component using exemplar models and the respiratory component by a subject-specific model (S4).
- The use of exemplar models improves the lowest error achieved by PCA model by 10% (compare experiments S3 and S4).

Table 1. Statistics of the motion amplitude of the liver points ($\|\Delta p_t^i\|$) for all subjects. The average value of mean, standard deviation (std) and maximum (max) over all subjects was 4.21 mm, 2.39 mm and 17.12 mm, respectively.

Subj	Amplitude of motion mean \pm std (max)(mm)	Subj	Amplitude of motion mean \pm std (max)(mm)	Subj	Amplitude of motion mean \pm std (max)(mm)
1	5.76 \pm 3.47 (20.93)	5	3.40 \pm 2.05 (19.49)	9	3.91 \pm 2.09 (14.29)
2	3.10 \pm 1.73 (17.26)	6	5.96 \pm 3.42 (19.71)	10	7.28 \pm 3.61 (21.15)
3	3.13 \pm 1.65 (11.51)	7	2.76 \pm 1.71 (11.21)	11	3.58 \pm 2.13 (16.75)
4	3.25 \pm 2.09 (16.68)	8	3.99 \pm 2.28 (18.40)	12	4.37 \pm 2.42 (18.02)

Table 2. Performance of the models without separating drift and respiratory motions in leave-one-out tests

	Generic Motion Models	Mean over all subjects j of		
		mE_j (mm)	sE_j (mm)	xE_j (mm)
G1	Subject-Specific Model	1.18	0.99	8.45
G2	PCA Population Model	1.08	0.82	8.12
G3	Exemplar Population Model	1.02	0.70	8.07

Table 3. Performance of separated drift and respiratory models in leave-one-out tests

	Separated Motion Models		Mean over all subjects j of		
	Drift Model	Respiration Model	mE_j (mm)	sE_j (mm)	xE_j (mm)
S1	PCA Population	PCA Population	1.16	0.84	9.02
S2	Exemplar Population	Exemplar Population	1.05	0.79	8.26
S3	PCA Population	Subject-Specific	0.97	0.77	8.65
S4	Exemplar Population	Subject-Specific	0.87	0.67	7.99

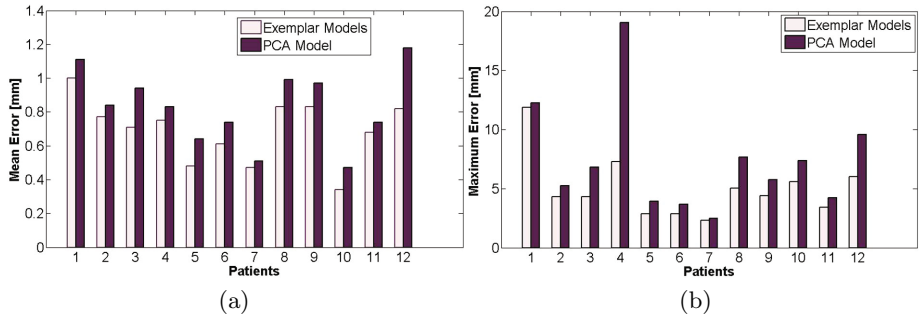


Fig. 2. Comparison of the performance of PCA and exemplar drift population models, showing (a) the mean and (b) the maximum error in predicting the drift of each point in the liver of the left-out subjects

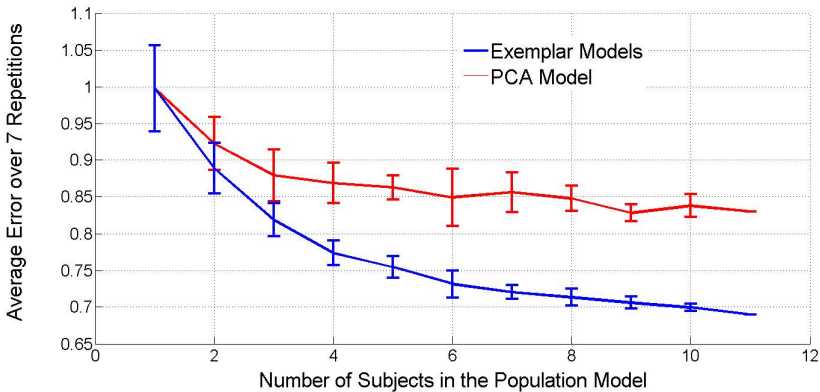


Fig. 3. Mean error of drift prediction over all subjects as a function of the size of the population using PCA and exemplar population models. The error bars show the standard deviation of the mean error from the 7 repetitions.

Fig. 2 shows a comparison of the drift prediction in experiments S3 and S4 for the individual subjects. It can be observed that the exemplar method provided a lower mean and maximum error than the PCA method for each subjects. The difference in mean prediction error of S4 and S3 was also statistically significant ($p < 0.001$). No other significance tests were performed.

The results of the learning capacity experiments are depicted in Fig. 3. The PCA model shows no substantial improvement when incorporating more than 5 subjects. In contrast, the exemplar model continues to improve as the number of considered subjects increases.

5 Discussion

Using Kullback-Leibler(KL) divergence as a distance measure between distributions, we observed that subjects 12 and 7 have the largest and the smallest sum of distances from the other subjects in the population. Also, the highest and lowest amounts of improvement using exemplar models are associated with subjects 12 and 7 respectively. This can be due to the fact that by fitting one PCA model to a population, the model captures the average motion pattern, performing poorly on uncommon subjects. The exemplar model is, however, more robust towards such outliers in motion patterns.

The most important advantage of exemplar models is their learning capacity. The trend seen in Fig. 3 suggests that exemplar models are more capable of using complementary information from a new subject and incorporating it into their models.

Furthermore, the exemplar method is incremental in the sense that adding new subjects to the model can be performed without re-calculating already considered cases. This attribute reduces the computational time and, together with the improved learning capacity of the approach, allows for step-by-step incorporation of cases where insufficient prediction is observed.

6 Conclusion

In this study, we addressed the problem of modeling and predicting two independent types of liver motion namely respiratory and drift motion in the presence of different breathing patterns among subjects. We proposed a method to generate an adaptable population model based on 4D motion fields obtained from 4DMRI images of 12 subjects.

Exemplar models were built for classes of the population represented by the individual subjects. Using a linear combination of the prediction of these exemplar models, the respiratory motion of a new subject was predicted, such that the most similar exemplar models have the highest influence.

We have explored different modeling methods for respiration and drift motion using PCA and exemplar models and concluded that a subject-specific PCA respiratory motion model based on pre-therapy data, combined with an exemplar population model for the drift achieve the lowest mean error. Finally, it was shown that the exemplar models have a higher learning capacity.

Acknowledgements. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°270186.

References

1. Tanner, C., Boye, D., Samei, G., Székely, G.: Review on 4D models for organ motion compensation. *Critical Reviews in Biomedical Engineering* 40(2), 135–154 (2012)
2. Blackall, J., Ahmad, S., Miquel, M., McClelland, J., Landau, D., Hawkes, D.: MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Phys. Med. Biol.* 51, 4147–4169 (2006)
3. Zhang, Q., Pevsner, A., Hertanto, A., Hu, Y., Rosenzweig, K., Ling, C., Mageras, G.: A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Med. Phys.* 34, 4772–4781 (2007)
4. Liu, X., Davis, B., Niethammer, M., Pizer, S., Mageras, G.: Prediction-driven respiratory motion atlas formation for 4D image-guided radiation therapy in lung. In: MICCAI, Int. Workshop on Pulmonary Image Analysis, pp. 71–82 (2010)
5. von Siebenthal, M., Székely, G., Lomax, A., Cattin, P.: Inter-subject Modelling of Liver Deformation During Radiation Therapy. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 659–666. Springer, Heidelberg (2007)
6. He, T., Xue, Z., Xie, W., Wong, S.T.C.: Online 4-D CT Estimation for Patient-Specific Respiratory Motion Based on Real-Time Breathing Signals. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part III. LNCS, vol. 6363, pp. 392–399. Springer, Heidelberg (2010)
7. Nguyen, T., Moseley, J., Dawson, L., Jaffray, D., Brock, K.: Adapting liver motion models using a navigator channel technique. *Med. Phys.* 36, 1061 (2009)
8. Klinder, T., Lorenz, C., Ostermann, J.: Prediction Framework for Statistical Respiratory Motion Modeling. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part III. LNCS, vol. 6363, pp. 327–334. Springer, Heidelberg (2010)
9. Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Handels, H.: Statistical modeling of 4D respiratory lung motion using diffeomorphic image registration. *IEEE T. Med. Imaging* 30(2), 251–265 (2011)
10. Preiswerk, F., Arnold, P., Fasel, B., Cattin, P.C.: A Bayesian Framework for Estimating Respiratory Liver Motion from Sparse Measurements. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) Abdominal Imaging. LNCS, vol. 7029, pp. 207–214. Springer, Heidelberg (2012)
11. Arnold, P., Preiswerk, F., Fasel, B., Salomir, R., Scheffler, K., Cattin, P.C.: 3D Organ Motion Prediction for MR-Guided High Intensity Focused Ultrasound. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part II. LNCS, vol. 6892, pp. 623–630. Springer, Heidelberg (2011)
12. Ramus, L., Commowick, O., Malandain, G.: Construction of Patient Specific Atlases from Locally Most Similar Anatomical Pieces. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part III. LNCS, vol. 6363, pp. 155–162. Springer, Heidelberg (2010)
13. Li, R., Lewis, J., Jia, X., Zhao, T., Liu, W., Wuenschel, S., Lamb, J., Yang, D., Low, D., Jiang, S.: On a PCA-based lung motion model. *Phys. Med. Biol.* 56, 6009 (2011)

14. Aha, D., Kibler, D., Albert, M.: Instance-based learning algorithms. *Machine Learning* 6(1), 37–66 (1991)
15. Chum, O., Zisserman, A.: An exemplar model for learning object classes. In: *CVPR*, pp. 1–8. IEEE (2007)
16. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. *IEEE Trans. PAMI* 24(4), 509–522 (2002)
17. Dudani, S.: The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man and Cybernetics* (4), 325–327 (1976)
18. Ahrendt, P.: The multivariate Gaussian probability distribution. Technical report, Technical University of Denmark (2005)
19. Efron, B., Tibshirani, R.: An introduction to the bootstrap. *Monographs on Statistics and Applied Probability*, vol. 57. Chapman & Hall/CRC (1993)

Semi-automatic Discrimination of Normal Tissue and Liver Cancer Lesions in Contrast Enhanced X-Ray CT-Scans*

Sanat Upadhyay¹, Manos Papadakis¹, Saurabh Jain², Gregory Gladish M.D.³, Ioannis A. Kakadiaris⁴, and Robert Azencott¹

¹ Department of Mathematics, University of Houston, TX 77204-3008, USA
mpapadak@math.uh.edu

² Center for Imaging Sciences, John Hopkins University, MD, USA

³ MD Anderson Cancer Center, The University of Texas, TX, USA

⁴ Department of Computer Science, University of Houston, TX, USA

Abstract. In this paper we present a set of 3D-rigid motion invariant texture features. We experimentally establish that when they are combined with mean attenuation intensity differences the new augmented features are capable of discriminating normal from abnormal liver tissue in arterial phase contrast enhanced X-ray CT-scans with high sensitivity and specificity. To extract these features CT-scans are processed in their native dimensionality. We experimentally observe that the 3D-rotational invariance of the proposed features improves the clustering of the feature vectors extracted from normal liver tissue samples.

Keywords: Liver cancer, 3D-texture classification, rotationally invariant features, soft tissue discrimination.

1 Introduction and Previous Work

The early detection of liver cancer lesions can potentially improve the management of various forms of liver cancer. Typically, liver lesions are identified using contrast enhanced CT scans acquired at different phases of perfusion of the hepatic parenchyma by the infused contrast agent. The task of identifying the lesions is performed by a radiologist using a large number of images generated from this multiphase CT acquisition and requires significant time and effort.

In this work, we present an algorithm and experimental results that demonstrate the feasibility of the development of a semi-automatic screening tool capable of detecting liver abnormalities in contrast enhanced x-ray CT-scans. Specifically, utilizing ideas proposed by Jain *et. al.* [4] we develop 3D-rigid motion invariant texture features. We experimentally establish that when these features are combined with mean attenuation intensity differences the new augmented features are capable of discriminating normal from abnormal liver tissue in arterial phase contrast enhanced X-ray CT-scans with high sensitivity and specificity.

* This work was partially supported by NSF DMS 0915242 award.

When scans are acquired during different perfusion phases of the contrast agent, liver lesions result in hypodense or hyperdense Regions of Interest (ROI) relative to normal hepatic parenchyma. Quite often, in the arterial phase hypodense ROIs are adjacent to hyperdense ones due to the increased vascularization of active cancerous lesions. The driving assumption in our approach is that liver tissue ROIs can be represented in a contrast-enhanced x-ray CT scan by two components, *3D-texture* and *local mean intensities*. The first of the two components captures the structure while the second provides the average intensity of the ROI, which is a traditional feature for tissue discrimination and is much more observable by the eye of the trained beholder than the former. We demonstrate though that local attenuation intensity averages are not by themselves robust enough to discriminate normal from abnormal tissue. Our results reveal that features capturing the structural characteristics of the 3D-textures associated with these tissue types in general perform better than the former, or at least equally well.

Texture-based lesion segmentation has been successfully used in the past for the detection of cancerous hepatic lesions [11]. In contrast to the herein proposed method, which is natively implemented in 3D, proposed texture features in the existing literature are extracted in a slice by slice manner by combining first and second order moments [3]. Very similar approaches have also been used to segment the liver from neighboring organs [7][15][12]. Apart from the fact that our methods are natively designed to work in 3D, a fundamental difference between previous texture-based approaches and our work is that they use significantly more complex classifiers. Liver segmentation and detection of cancerous lesions has been mostly performed with non-texture based methods as in [1][9][8][17][14][10] where the difference in attenuation intensity between more and less contrast-perfused ROIs is used for feature extraction or as in [5][2] where deformable models are utilized to generate the boundaries between normal and abnormal tissues. However, both of these approaches mostly limit the detection of cancer lesions to the hypodense ones, because differences in average intensities are typically the dominant discriminative features.

In spite of the significant successes in the field of hepatic tumor detection and segmentation, our work opens an unexplored direction. The novelty of our approach relies first, on the use of 3D isotropic multiscale analysis for the extraction of 3D-rigid motion invariant texture features; second, on augmenting these texture features with attenuation intensity-based features. The proposed 3D-rigid motion invariant features allow feature vectors from normal tissue samples to form clusters that are more well-defined than the clusters formed when the 3D-texture features do not account for 3D-rotations. This enhances the discriminatory power of the proposed features (Fig. 3). Our experimental results are not directly comparable with the results of others, because we only test the discriminative power of our features on sets of ROIs and we don't segment normal from abnormal liver tissue. However, it appears that the proposed features can be used for tissue discrimination with high sensitivity and specificity rates rendering them as a promising tool for developing segmentation algorithms.

2 Methods

Our tissue classification scheme consists of two levels of classifiers. The design of both levels is traditional. The first level consists of an ensemble of SVMs classifiers which at the second level decide by majority voting whether a tissue ROI is normal or abnormal. The SVM classifiers use low-dimensional feature vectors whose components express the statistical disparity at one or more scales of the 3D-texture corresponding to a given liver tissue ROI for the 3D-texture of a normal reference ROI and the difference of average intensities between the two ROIs. To develop these classifiers the human operator selects a small number of reference normal ROIs from an x-ray CT-scan that is examined. The proposed feature design takes into account that a liver consists of soft tissues with varying 3D-orientations thus requiring features to be invariant under 3D-rotations and translations. In particular, cancer will tend to develop along blood vessels, which themselves appear with a varying degree of 3D-orientations. Moreover, malignancies form their vasculature with an even richer orientation variation in 3D. Attenuation intensity local average-based features would automatically be invariant to 3D-rigid motions, and therefore insensitive to the variety of 3D-orientations of the patterns formed by tissues of interest, but 3D-texture features must be specifically designed to be 3D-rigid motion invariant. We discuss the details of this design in the next paragraph.

2.1 3D-Texture Based Features

To discriminate 3D-textures corresponding to soft-tissue ROIs we assume that both ROIs have zero mean. The texture component of the proposed features is derived by combining 3D-rigid motion invariant monoscale ‘distances’ between texture signatures derived by fitting order-one Gaussian Markov Random Field (GMRF) models to the orbits of 3D-texture rotations corresponding to the zero-mean tissue ROIs, as proposed in [4,16]. In this manner, our features quantify 3D-texture disparities at various scales. A 3D-texture \mathbf{X} is modeled as a spatially homogeneous random field defined on the continuous domain \mathbb{R}^3 . Image acquisition generates the realization of \mathbf{X} in the form of a digital 3D-image whose values at the points of a discrete sampling lattice Λ are the exact same values of its ‘continuous parent’ at the grid points of Λ . Hence, an α -rotation of the discrete texture \mathbf{X} is the restriction of the α -rotation of its continuous parent on Λ . So, rotating the discrete texture \mathbf{X} amounts to rotating the autocovariance ρ of its continuous parent.

Using the approximation of ρ by the empirical autocovariance function ρ_0 we ‘fit’ a very simple order-one GMRF model to the data $\rho_0(\mathbf{k})_{\mathbf{k} \in \Lambda}$ from which the autocovariance matrix of the fitted model $\hat{\Sigma}$ and its inverse $(\hat{\Sigma})^{-1}$ are computed following a method originally proposed in [6] adopted for 3D in [4,13]. By taking any α -rotation of ρ_0 and fitting the same model to the rotated ρ_0 $(\hat{\Sigma})\alpha$ is obtained. Since, 3D-shifts induced by the action of Λ do not affect the discrete texture \mathbf{X} , due to the spatial homogeneity assumption it follows that the orbit $\Gamma_{\mathbf{X}}(\alpha) = (\hat{\Sigma})\alpha$, $\alpha \in SO(3)$ is the *monoscale 3D-rigid motion covariant texture*

signature of the observed texture \mathbf{X} at the scale corresponding to the density of the lattice Λ . This led Jain *et al* [4] to propose a 3D-rigid motion invariant texture ‘distance’ at this scale

$$Rdist(\mathbf{X}, \mathbf{Y}) := \min_{\tau \in SO(3)} \int_{SO(3)} \text{KLdist}(\Gamma_{\mathbf{X}}(\alpha), \Gamma_{\mathbf{Y}}(\tau\alpha)) d\alpha, \quad (1)$$

where $\text{KLdist}(\Sigma_1, \Sigma_2) = \frac{1}{2} \text{Trace}(\Sigma_2^{-1}\Sigma_1 + \Sigma_1^{-1}\Sigma_2 - 2I_{N \times N})$. We stress that the 3D-texture corresponding to a tissue ROI is almost never an order one GMRF. Yet, we carry out our computations as if it were such a GMRF. We use this computationally simple and numerically efficient stochastic model as a probe for tissue discrimination and not as a texture model for soft tissues. The 3D-texture features for the promised liver tissue SVM classifiers are constructed via [11].

2.2 Feature Extraction

Let $\{\mathbf{s}_k\}_{k=1,2,\dots,N}$ be 3D volumes from normal or abnormal liver tissues and $\mathbf{\Lambda}_k := \mathbf{s}_k \cap \Lambda$. Now, fix a sample \mathbf{n} which is known to be normal to be the reference normal. For each one of the $\{\mathbf{s}_k\}_{k=1,2,\dots,N}$ we derive the feature vector $\mathbf{f}(\mathbf{s}_k; \mathbf{n})$ relative to \mathbf{n} according to the following algorithm.

The first component of the feature vector $\mathbf{f}(\mathbf{s}_k; \mathbf{n})$ is

$$D_M(\mathbf{s}_k, \mathbf{n}) := \frac{|\bar{\mathbf{s}}_k - \bar{\mathbf{n}}|}{\sqrt{\text{var}(\mathbf{s}_k) + \text{var}(\mathbf{n})}}. \quad (2)$$

D_M standardizes the statistical disparity due to the difference in the average attenuation intensity between \mathbf{s}_k and \mathbf{n} . To form the remaining components of the feature vector $\mathbf{f}(\mathbf{s}_k; \mathbf{n})$, we use the 3D-rigid motion invariant statistical disparity $RD_j(\mathbf{s}_k, \mathbf{n})$ at scales $j = 0, -1$, between the 3D textures corresponding to \mathbf{s}_k and \mathbf{n} where

$$RD_j(\mathbf{s}_k, \mathbf{n}) = \max \left\{ 0, \frac{Rdist_j(\mathbf{s}_k, \mathbf{n}) - \text{diam}_{\mathbf{n}}(j)}{\text{diam}_{\mathbf{n}}(j)} \right\} \quad (3)$$

[4]. If for a given sample \mathbf{s}_k and at some scale j , $RD_j(\mathbf{s}_k, \mathbf{n})$ is large, then we can conclude that the sample \mathbf{s}_k has a different 3D texture than the reference normal \mathbf{n} \mathbf{s}_k thus more the tissue from which \mathbf{s}_k originated is likely to be abnormal. To compute $RD_j(\mathbf{s}_k, \mathbf{n})$ we make some non-trivial modifications to the algorithm proposed in [16] to make it applicable to ROIs that are not 3D cubes. We describe those modifications below:

Computation of $RD_j(\mathbf{s}_k, \mathbf{n})$: (i) Adjust intensity values in both \mathbf{s}_k and \mathbf{n} to have zero mean.

(ii) Upsample each \mathbf{s}_k to a twice the denser grid as in [16]. The upsampled 3D-texture sample is convolved with the isotropic low-pass synthesis filter H_0 :

$$H_0(\xi) = \begin{cases} 1 & |\xi| \leq \frac{1-\beta}{2\tau}, \\ \frac{1}{2} \left[1 + \cos \left(\frac{\pi\tau}{\beta} \left(|\xi| - \frac{1-\beta}{2\tau} \right) \right) \right] \frac{1-\beta}{2\tau} & \frac{1-\beta}{2\tau} < |\xi| \leq \frac{1+\beta}{2\tau}, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where $\tau = 100/84$ and $\beta = 1/7$. The isotropy of H_0 increases the accuracy of the computation of ρ_0 .

(iii) Let \mathbf{t} be a node in \mathbf{s}_k , then a *neighborhood* of \mathbf{t} is $\eta_{\mathbf{t}} := (\mathbf{t} + W) \subset \mathbf{\Lambda}_k$ where W is a symmetric neighborhood of the origin. We set $W = W^+ \cup W^-$ where $W^- = -W^+$ and $W^+ = \{(2^{-j+1}, 0, 0), (0, 2^{-j+1}, 0), (0, 0, 2^{-j+1})\}$. The order-one GMRF model limits interactions within W . Also define $\mathbf{\Lambda}'_k \subset \mathbf{\Lambda}_k$ to satisfy $(\eta_{\mathbf{t}} + W) \subset \mathbf{\Lambda}_k$ for every $\mathbf{t} \in \mathbf{\Lambda}'_k$. We extract our statistics from $\mathbf{s}_k|_{\mathbf{\Lambda}'_k}$.

(iv) Compute the empirical auto-covariance matrix ρ_0 of $\mathbf{s}_k|_{\mathbf{\Lambda}'_k}$ via:

$$\rho_0(\mathbf{t}) = \frac{1}{|\mathbf{\Lambda}'_k|} \sum_{\mathbf{r} \in \mathbf{\Lambda}'_k} s_{k_r} s_{k_{\mathbf{r}+\mathbf{t}}}, \quad \text{for all } \mathbf{t} \text{ such that } \|\mathbf{t}\|_{\infty} \leq 2^{-j+2}.$$

where $|\mathbf{\Lambda}'_k|$ denotes the number of voxels in $\mathbf{\Lambda}'_k$; ρ_0 is of size $(2^{-j+3} + 1)^3$.

(v) Any of the 3D textures need not satisfy $\rho_0(\mathbf{t}) = \rho_0(-\mathbf{t})$. So we artificially symmetrize ρ'_0 by setting $\rho'_0(\mathbf{t}) := \frac{1}{2}[\rho_0(\mathbf{t}) + \rho_0(-\mathbf{t})]$ for all $\mathbf{t} \in \mathbf{\Lambda}'_k$ such that $\|\mathbf{t}\|_{\infty} \leq 2^{-j+2}$. To simplify the notation from now on we use $\rho'_0 = \rho_0$.

(vi) Let $\mathbf{y}_{\mathbf{r}} = [s_{k_{\mathbf{l}}} + s_{k_{-\mathbf{l}}}], \mathbf{l} \in (\mathbf{r} + W^+)$. Define $\mathbf{Y} = [\mathbf{y}_{\mathbf{r}}], \mathbf{r} \in \mathbf{\Lambda}'_k$. The least squares estimates $\hat{\sigma}$ and $\hat{\theta}$ of the order-one GMRF model that fits the data are given by the statistics: $\hat{\theta}(\mathbf{s}_k) = (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{s}_k$ and $\hat{\sigma}^2(\mathbf{s}_k) = \frac{1}{|\mathbf{\Lambda}'_k|} \left(\mathbf{s}_k^T \mathbf{s}_k - \hat{\theta}^T \mathbf{Y}^T \mathbf{s}_k \right)$

(vii) $Y^T \mathbf{s}_k$ and $Y^T Y$ are given by $(\mathbf{Y}^T \mathbf{s}_k)_{\mathbf{r}} = |\mathbf{\Lambda}'_k|(\rho_0(\mathbf{r}) + \rho_0(-\mathbf{r})) \quad \forall \mathbf{r} \in W^+$

$(\mathbf{Y}^T \mathbf{Y})_{(\mathbf{t}, \mathbf{r})} = |\mathbf{\Lambda}'_k|[\rho_0(\mathbf{r}-\mathbf{t}) + \rho_0(\mathbf{r}+\mathbf{t}) + \rho_0(-\mathbf{r}-\mathbf{t}) + \rho_0(-\mathbf{r}+\mathbf{t})] \quad (\mathbf{t}, \mathbf{r}) \in W^+ \times W^+.$

(viii) The computation of $(Y^T Y)^{-1}$ is performed as in [16].

By iterating for a finite set of rotations the previous steps the rotationally covariant signatures $\Gamma_{\mathbf{s}_k}$ and $\Gamma_{\mathbf{n}}$ are generated as in [4] and finally $RD_j(\mathbf{s}_k, \mathbf{n})$ is computed for $j = 0, -1$ using the isotropic low pass filter [4] with $\tau = \frac{2^{-j} \cdot 100}{84}$. We keep the low-pass output at the original resolution. To extract the order-one GMRF statistics we use the interactions of pixels 2^{-j+1} apart and we repeat all of the previous steps with the exception of step (ii). As noted in [4] observing textures at lower scales compensates for the loss of the tissue textures information due the use of the order-one GMRF model.

3 Experiments and Results

Arterial phase X-ray CT scans of liver from six patients were obtained. All scans were obtained with almost cubic voxels which is necessary for our method. From each of the six CT-scan series, we selected a set of eight normal and twelve abnormal ROIs, chosen by an expert radiologist. The abnormal ROIs contained one or more cancer lesions at different stages of maturity and size. For each of the series we perform three experiments. First, we use each one of the normal ROIs \mathbf{n} , as a *reference normal*. Then, we compute three sets of feature vectors. Using the feature vectors we develop a two class SVM-classifier corresponding to \mathbf{n} with each one of these sets of features. The kernel of the SVM-classifier is

approximated by Radial Basis functions and for our implementation we use the free package LibSVM. Next we define the three said types of features.

(a) $\mathbf{f}_1(\mathbf{s}_k) = (D_M(\mathbf{s}_k, \mathbf{n}))$, classifies samples \mathbf{s}_k by exclusively using differences of ROI mean attenuation intensities.

(b) $\mathbf{f}_2(\mathbf{s}_k) = (RD_0(\mathbf{s}_k, \mathbf{n}), RD_{-1}(\mathbf{s}_k, \mathbf{n}))$, classifies samples \mathbf{s}_k using their 3D-texture disparities at the original scale and at one scale coarser than the original. 3D-displacements and 3D-rotations of ROIs do not influence the statistical disparity of the textures of any two ROIs. Any directional characteristics native to each texture contribute to this texture disparity feature only with respect to their relative 3D-orientations.

(c) $\mathbf{f}_3(\mathbf{s}_k) = (RD_0(\mathbf{s}_k, \mathbf{n}), RD_{-1}(\mathbf{s}_k, \mathbf{n}), D_M(\mathbf{s}_k, \mathbf{n}))$ combines multiscale texture disparity with differences in average attenuation intensity.

To validate the performance of each set of the features above, we test each of the SVM-classifiers associated with a normal tissue ROI using only this set of features. We also test a classifier combining input from each one of these classifiers by majority voting referred to as *SVM-voting* separately for features \mathbf{f}_1 , \mathbf{f}_2 and \mathbf{f}_3 . We estimate the accuracy of these classifiers in two ways.

Experiment 1: We randomly divide 20 samples in two groups, with 4 normal and 6 abnormal ROIs in each group. We perform a *two fold cross validation* by alternately training the SVM classifier on each group and testing it on the other. The average accuracy of each pair of classifiers is referred to as the accuracy of the two-fold classification. We repeat the previous test 100 times. The average accuracy of the two-fold classification from these 100 random trials is shown in Tables 1, 2. The average accuracy of the SVM-voting classifier for the same 100 random trials is also shown in Tables 1 and 2.

Experiment 2: This leave-two-out experiment essentially assesses the generalization capacity of the proposed classifiers. From the sample of twenty tissues from each patient we use 18 of them for training and two for testing. Leave-two-out accuracy is calculated for each of the eight SVM classifiers corresponding to each of the normal ROIs for each of the CT-scan series and for each of the proposed features \mathbf{f}_1 , \mathbf{f}_2 and \mathbf{f}_3 . We also assess the performance of the SVM-voting classifier. In the same way Tables 1 and 2 provide the average performance of

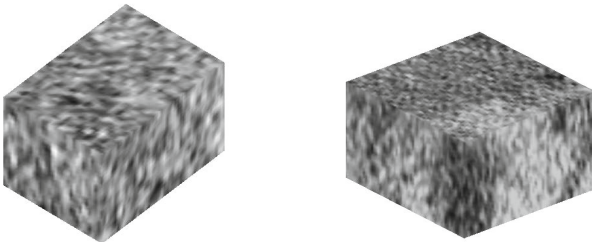


Fig. 1. Typical 3D-view of the texture of a normal liver ROI (left) and of an abnormal (neoplastic) liver ROI (right)

Table 1. The entries of the table denotes the accuracy in percentage for each set of proposed features. SVM Classifiers are defined for each set of features and relative to each of the normal ROIs in each CT-scan series of Livers from 2 patients (L1, L2). Then we observe the average accuracy obtained from N1 to N8. *SVM-Voting* gives the classification based on the majority voting of *reference normals* used for training.

	Reference Normal	Two fold cross-validation			Leave-two-out-validation		
		Features used			Features used		
		D_M	(RD_0, RD_{-1})	(RD_0, RD_{-1}, D_M)	D_M	(RD_0, RD_{-1})	(RD_0, RD_{-1}, D_M)
L1	N1	95.0	94.4	100	92.7	80.7	100
	N2	75.8	90.4	97.7	77.6	84.4	100
	N3	77.0	91.6	98.1	71.4	90.1	100
	N4	95.6	95.1	97.8	100	91.1	100
	N5	94.9	95.1	98.1	99.5	93.8	100
	N6	94.9	94.4	99.2	99.5	89.6	100
	N7	95.5	92.8	99.1	100	90.6	100
	N8	95.2	95.0	97.9	100	95.8	100
	Average	90.5	93.6	98.5	92.6	89.5	100
	<i>SVM-Voting</i>	94.8	94.0	98.6	100	89.1	100
L2	N1	90.3	91.0	99.2	78.1	79.2	100
	N2	92.3	79.5	99.4	95.8	95.8	100
	N3	91.1	91.3	99.7	86.5	87.0	100
	N4	92.0	88.6	99.4	90.6	86.5	100
	N5	92.2	89.0	98.5	75.0	81.8	100
	N6	92.3	84.5	96.4	81.3	92.2	99.5
	N7	85.4	93.5	99.6	69.3	91.1	100
	N8	88.9	91.2	98.4	88.0	87.0	100
	Average	90.5	88.6	98.8	83.1	87.6	99.9
	<i>SVM-Voting</i>	91.2	90.2	99.7	80.2	87.0	100

these classifiers tested on all possible 96 partitions of the ROI set from each CT-scan series.

4 Discussion

In this paper we present a novel set of features combining information of multi-scale texture disparity with the difference between average attenuation values for a given pair of texture patches one of which corresponds to normal tissue. These features exploit the full power of the 3D information modern scanners provide. We develop our features and the associated SVM classifiers using normal tissue reference ROIs only, because this type of tissue is less diverse within the same organ. Since the proposed disparities between normal tissues are also typically relatively small, because not absolute but only relative 3D-orientations of directional characteristics influence the evaluation of textural disparities (Fig. 3) making disparity assessments obtained from different SVM classifiers to agree. Tables 1, 2 support this claim by manifesting the robust assessment of textural similarity between normal tissues and the dissimilarity between normal and abnormal ones enabling thus a highly accurate SVM voting classification.

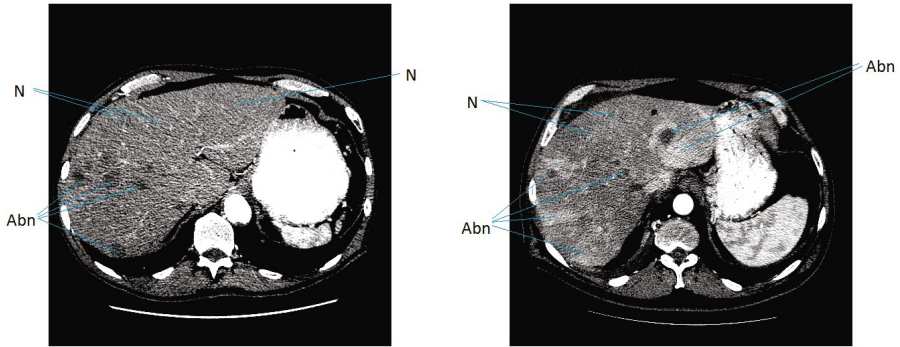


Fig. 2. Examples of our choices of normal (N) and diseased (Abn) regions, whose cross-sections are shown on 2D slices. We selected many different kind of abnormalities for our experiments, which includes tumors of different sizes and from different stages.

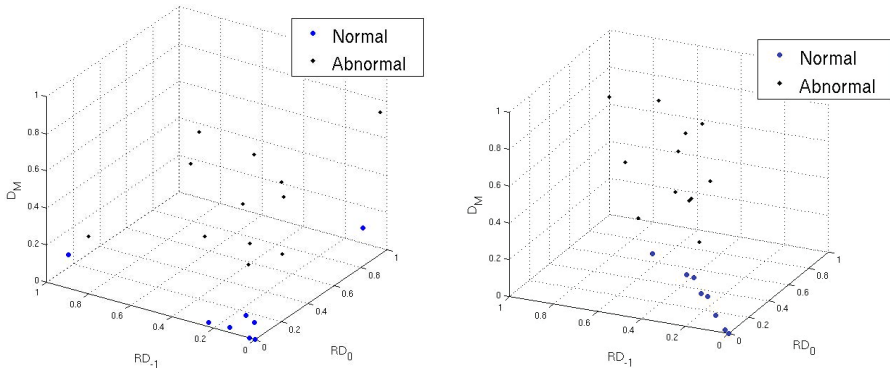


Fig. 3. Scatterplots show that the use of 3D-rotationally invariance texture disparities improves normal tissue ROI feature vector clustering. Left: RD_j 's given by texture distance $Rdist_j$ (Eq. (II)). Most feature vectors from normal tissue ROIs cluster around the origin and clearly away from their abnormal ROI counterparts. Right: Texture disparities are computed using $KLdist$ only, without averaging over 3D rotations. The normal tissue ROI cluster is less pronounced and it is more proximal to abnormal tissue ROI feature vectors.

Table 2. Continuation of Table 1 to show the classification accuracy obtained from the two experiments for rest of the four patients (L3-L6)

	Reference Normal	Two fold cross-validation			Leave two out validation		
		Features used			Features used		
		D_M	(RD_0, RD_{-1})	(RD_0, RD_{-1}, D_M)	D_M	(RD_0, RD_{-1})	(RD_0, RD_{-1}, D_M)
L3	N1	73.4	83.4	86.2	67.2	58.3	90.6
	N2	73.6	74.4	71.1	83.3	78.1	49.5
	N3	75.0	82.9	84.7	75.5	78.1	93.2
	N4	69.4	78.2	73.4	49.5	65.6	62.5
	N5	77.0	79.4	79.7	92.2	66.7	93.8
	N6	74.6	81.6	79.7	72.9	67.2	93.8
	N7	76.5	84.2	85.6	97.4	77.1	81.3
	N8	75.1	85.2	82.9	99.5	95.3	85.4
	Average	74.3	81.1	80.4	79.7	73.3	81.3
	<i>SVM-Voting</i>	74.4	83.2	83.0	91.1	70.3	91.1
L4	N1	93.1	71.4	84.9	83.3	56.3	71.4
	N2	89.0	70.4	89.1	73.4	63.0	92.7
	N3	95.2	68.6	90.6	84.4	52.6	87.5
	N4	92.5	70.1	93.0	79.2	76.6	95.8
	N5	89.5	78.0	91.3	83.3	85.4	90.1
	N6	95.0	68.8	90.5	93.8	72.9	92.7
	N7	95.2	74.4	91.2	92.7	69.3	94.8
	N8	95.3	80.0	91.6	97.4	87.5	93.8
	Average	93.1	72.7	90.3	85.9	70.4	89.8
	<i>SVM-Voting</i>	93.5	69.7	91.5	90.6	67.2	93.8
L5	N1	95.0	71.4	93.5	82.8	50.5	91.1
	N2	95.1	70.0	92.0	87.0	66.7	94.8
	N3	95.1	70.3	93.2	78.1	68.8	99.0
	N4	99.2	75.1	95.0	100	59.9	92.2
	N5	95.1	72.1	92.1	79.2	56.8	93.8
	N6	97.5	72.3	92.9	100	76.0	95.8
	N7	96.7	73.7	91.7	99.5	58.9	88.5
	N8	95.0	70.6	93.0	69.3	47.4	99.0
	Average	96.1	71.9	92.9	87.0	60.6	94.3
	<i>SVM-Voting</i>	95.0	72.3	93.2	91.1	55.7	95.8
L6	N1	96.2	91.1	99.6	99.5	100	100
	N2	95.1	74.2	96.8	96.9	81.3	100
	N3	97.9	74.0	98.1	100	80.2	100
	N4	98.3	81.3	98.9	100	75.5	100
	N5	98.4	66.7	95.0	100	80.7	99.5
	N6	95.2	82.2	97.6	96.4	85.9	100
	N7	97.9	81.6	99.9	100	81.3	100
	N8	95.0	70.7	96.4	93.8	51.6	100
	Average	96.7	77.7	97.8	98.3	79.6	99.9
	<i>SVM-Voting</i>	95.5	79.8	98.6	100	82.8	100

References

1. Abdel-massieh, N.H., Hadhoud, M.M., Amin, K.M.: Automatic liver tumor segmentation from ct scans with knowledge-based constraints. In: 5th Cairo International Biomedical Engineering Conference, pp. 215–218. IEEE (December 2010)
2. Chen, T., Metaxas, D.: A hybrid framework for 3D medical image segmentation. *Journal of Medical Image Analysis* 9, 547–565 (2005)

3. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics SMC-3*(6), 610–621 (1973)
4. Jain, S., Papadakis, M., Upadhyay, S., Azencott, R.: Rigid-motion-invariant classification of 3-D textures. *IEEE Transactions on Image Processing* 21(5), 2449 (2012)
5. Jolly, M.P., Grady, L.: 3D general lesion segmentation in CT. In: 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE (2008)
6. Kashyap, R.L., Chellappa, R.: Estimation and choice of neighbors in spatial-interaction models of images. *IEEE Trans. Information Theory* 1, 60–72 (1983)
7. Liu, J., Hu, Q., Chen, Z., Heng, P.: Adaptive liver segmentation from multi-slice CT scans. In: Peng, Y., Weng, X., Magjarevic, R. (eds.) 7th Asian-Pacific Conference on Medical and Biological Engineering. IFMBE Proceedings, vol. 19, pp. 381–384. Springer, Heidelberg (2008)
8. Massoptier, L., Casciaro, S.: A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans, vol. 18, pp. 1658–1665. Springer, Heidelberg (2008)
9. Militzer, A., Hager, T., Jäger, F., Tietjen, C., Hornegger, J.: Automatic detection and segmentation of focal liver lesions in contrast enhanced CT images. In: 20th International Conference on Pattern Recognition, pp. 2524–2527 (August 2010)
10. Park, S.-J., Seo, K.-S., Park, J.-A.: Automatic Hepatic Tumor Segmentation Using Statistical Optimal Threshold. In: Sunderam, V.S., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (eds.) ICCS 2005. LNCS, vol. 3514, pp. 934–940. Springer, Heidelberg (2005)
11. Pescia, D., Paragios, N., Chemouny, S.: Automatic detection of liver tumors. In: 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 672–675 (2008)
12. Pham, M., Susomboon, R., Disney, T., Raicu, D., Furst, J.: A comparison of texture models for automatic liver segmentation. In: Pluim, J.P.W., Reinhardt, J.M. (eds.) *Proc. Medical Imaging 2007: Image Processing of SPIE*, vol. 6512 (2007)
13. Rangelova, E.B.: Segmentation of textured images on three-dimensional lattices. PhD Thesis, University of Dublin, Trinity College (2002)
14. Seo, K.S.: Automatic Hepatic Tumor Segmentation Using Composite Hypotheses. In: Kamel, M.S., Campilho, A.C. (eds.) ICIAR 2005. LNCS, vol. 3656, pp. 922–929. Springer, Heidelberg (2005)
15. Susomboon, R., Raicu, D., Furst, J., Johnson, T.B.: A co-occurrence texture semi-invariance to direction, distance and patient size. In: Reinhardt, J.M., Pluim, J.P.W. (eds.) *Proc. Medical Imaging 2008: Image Processing*, vol. 6914. SPIE (2008)
16. Upadhyay, S., Jain, S., Papadakis, M., Azencott, R.: 3D-rigid motion invariant discrimination and classification of 3D-textures. In: Papadakis, M., Van De Ville, D., Goyal, V. (eds.) *Proc. Wavelets and Sparsity XIV, Proceedings SPIE*, vol. 8138, p. 813821 (2011), <http://dx.doi.org/10.1117/12.891721>
17. Zhou, J., Xiong, W., Tian, Q., Qi, Y., Liu, J., Leow, W.K., Han, T., Venkatesh, S.K., Wang, S.C.: Semi-automatic segmentation of 3d liver tumors from CT scans using voxel classification and propagational learning. In: MIDAS (2008)

Evaluation of Diffusion Filters for 3D CTA Liver Vessel Enhancement

Ha Manh Luu, Adriaan Moelker, Camiel Klink, Adrienne Mendrik,
Wiro Niessen, and Theo van Walsum

Biomedical Imaging Group Rotterdam, Departments of Medical Informatics
and Radiology, Erasmus MC, University Medical Center Rotterdam,
Dr. Molewaterplein 50, 3015 GE Rotterdam, The Netherlands
haluumanh@erasmusmc.nl

Abstract. We present an evaluation of five diffusion filters for liver vessel enhancement in 3D CTA datasets of the liver. 3D CTA liver images are generally noisy, with limited contrast between vessels and parenchyma, especially for the small vessels. We investigate the performance of five (an)isotropic diffusion filters: Regular Perona-Malik, Coherence-Enhancing Diffusion, Edge-Enhancing Diffusion, Hybrid Diffusion with Continuous Switch and Vessel Enhancing Diffusion on a set of 14 abdominal CTA clinical datasets. The evaluation is based on signal to noise improvement. A parameter optimization is performed on 7 training images, after which the optimal versions of the filters are compared on 7 test images. The results demonstrate that all the diffusion filters improve SNR of the images and Hybrid Diffusion with Continuous Switch and Vessel Enhancing Diffusion gives the largest increase in SNR.

Keywords: Diffusion, HDCS, VED, CED, RPM, EED, 3D CTA, liver vessel.

1 Introduction

Liver vessel analysis is relevant for several clinical applications, a.o. for planning and guidance in minimally invasive interventions. Particularly, segmentation of the portal and hepatic veins is relevant for procedures such as liver surgery, TIPS and RFA. Arterial segmentation is relevant for e.g. chemo-embolization procedures in the liver.

Liver vessel segmentation is challenging because 3D abdominal CTA is noisy, has variable contrast between the vessel and liver parenchyma, and because of the complex topology and varying sizes of the vessels. The quality of the CTA images depends on radiation dose, amount of contrast agent, and timing of data acquisition with respect to dose injection. Image quality could be improved at the expense of increased radiation dose or increased contrast usage, but both strategies are clinically not acceptable.

Several liver vessel segmentation methods have been developed so far, a review of liver vessel segmentation can be found at [1]. Most of the works relating to liver vessel segmentation on 3D clinical CTA use filters to reduce the noise and enhance the vessel structure. Mainly, multi-scale Hessian based filters (Frangi, Sato, Erdt)

[2, 4, 5, 7] have been used in these studies. However, these studies lack an evaluation on to what extent preprocessing improves the segmentation results. A comparison of multi-scale Hessian based filters can be found at [3], but that study only demonstrates the effect of these Hessian-based filters without any quantitative measure. The study also shows that Hessian-based approaches may not work well at bifurcations and vessels with high curvature.

The purpose of our study is to quantitatively evaluate the effect of diffusion filters on vessel enhancement in 3D CTAs of the liver. We investigate both the optimal filter settings for these filters, and compare the optimal versions of each filter. For the filters, we choose five well-known (an)isotropic diffusion filters: Regular Perona-Malik (RPM), Contrast-Enhancing Diffusion (CED), Edge-Enhancing Diffusion (EED), Vessel-Enhancing Diffusion (VED) and Hybrid Diffusion with Continuous Switch (HDCS). RPM and EED have been applied to 3D rotational angiography images by Meijering et al [10]. They showed that EED works well in smoothing the vessel, while RPM can preserve small vessels. CED, introduced by Weickert [9], can filter tube-like structures. VED was published by Manniesing et al in 2006 [8]. This filter uses Hessian-based multi-scale filter's responses to adjust diffusion scheme. Finally, HDCS was published by Mendrik et al in 2009 [10]. This filter combines both advantages of EED and CED to filter both homogeneous areas and vessel structures.

2 Methodologies

All filters we evaluated are diffusion filters. The main idea of diffusion filters comes from PDE [8] $u_t = \text{div}(D \cdot \nabla u)$ where ∇u is the gradient of the image and D is the diffusion tensor, which steers the diffusion. If the diffusion tensor D is replaced by a scalar-valued diffusivity g the diffusion will be isotropic. Whereas RPM is an isotropic filter, that only changes the amount of smoothing based on local gradient magnitude, the other filters are anisotropic filters, that not only locally change the level, but also the direction of smoothing by adapting the diffusion tensor. Each of the five filters is described in more detail below.

RPM: Perona and Malik (1990) introduced an isotropic nonlinear diffusion as described by $u_t = \text{div}(g(|\nabla u|) \cdot \nabla u)$. The scalar-valued diffusivity $g(|\nabla u|)$ is function of the gradient $|\nabla u|$, causing filtering in homogenous areas while retaining edges with high gradient. Catta [10] proposed the following scalar-valued diffusivity function for the non-linear diffusion using Gaussian derivative at scale σ :

$$g(|\nabla u_\sigma|) = 1 - \frac{-C}{e^{(|\nabla u_\sigma|^2 / \lambda^2)^4}}, \quad (1)$$

where $C = 3.1488$ and λ is contrast parameter. The contrast parameter λ acts as threshold scale for gradient $|\nabla u_\sigma|$. If the gradient is large compared to the contrast parameter, i.e $|\nabla u_\sigma|^2 \gg \lambda^2$, this results in $g(|\nabla u_\sigma|) \approx 0$, reducing the amount of diffusion. Therefore strong edges, where the gradient of u is large, are preserved. Parameter σ is the scale of Gaussian gradient. Value of σ should be chosen based on noise variance and the size of the small structures we want to retain.

EED: Weickert et al (1997) included information of orientation in diffusion scheme. Instead of using a scalar diffusivity function, they used a diffusion tensor D which was constructed from the tensor product:

$$J_\rho(\nabla u_\sigma) = K_\rho * (\nabla u_\sigma \nabla u_\sigma^T). \quad (2)$$

J_ρ is a positive symmetric matrix. By eigendecomposing $J_\rho = V.M.V^T$, they can extract eigenvalues μ_i ($i = 1-3$, $\mu_1 > \mu_2 > \mu_3$), the diagonal elements of matrix M , and corresponding direction eigenvectors V_1 , V_2 and V_3 . Because eigenvalue μ_1 is the largest eigenvalue, eigenvectors V_1 is in the direction of highest contrast (edges). The diffusion tensor is defined as $D = V.\Lambda.V^T$ with eigenvectors V as the same eigenvectors of tensor product J_ρ . Eigenvalues of diffusion tensor of EED, Λ , are defined as: $\lambda_{e_2} = \lambda_{e_3} = 1$ and

$$\lambda_{e_1} = \begin{cases} 1 & |\nabla u_\sigma| = 0 \\ \frac{-c}{1 - e^{-(|\nabla u|^2/\lambda_e^2)^4}}, & |\nabla u_\sigma| > 0 \end{cases}, \quad (3)$$

where, similar to RPM, $C = 3.1488$, and λ_e is the contrast parameters. This diffusion tensor results in large isotropic diffusion in flat areas where the gradient is small, and performs almost no diffusion in the direction along which gradient is the highest (V_1). K_ρ in equation (2) acts as smoothing of gradient (Gaussian convolution with kernel ρ). If gradient ∇u_σ has large range of values, ρ should be high enough to smooth the product. Otherwise, ρ should be small enough to capture small changes in gradient.

CED: Weickert (1999) included a coherence factor in diffusion process. The coherent factor is defined as:

$$\kappa = (\mu_1 - \mu_2)^2 + (\mu_2 - \mu_3)^2 + (\mu_3 - \mu_1)^2, \quad (4)$$

This factor measures the relation of each pair of the eigenvalues. If a structure is tubular, V_1 , V_2 direction are in the direction of high contrast and V_3 in the direction of little contrast. Thus in the case of tubular structures, $\mu_1 \approx \mu_2 \gg \mu_3$, and thus κ has high a value. Otherwise, κ has small value. Eigenvalues of diffusion tensor of CED are defined as: $\lambda_{c_1} = \lambda_{c_2} = \alpha$ and

$$\lambda_{c_3} = \begin{cases} \alpha & \kappa = 0 \\ \alpha + (1 - \alpha)\exp(-C/\kappa), & \kappa > 0 \end{cases}, \quad (5)$$

where $C = 3.1488$ and α is very small. For tube-like structures, κ is large, the diffusion mainly occurs in the direction V_3 (least contrast). Therefore CED only blurs along tubular structures. For plate-like structures, κ is small, resulting in small isotropic diffusion (depends on α).

HDCS: Mendrik (2009) introduced HDCS as a combination of CED and EED. The main idea is that to use a voting criterion to decide whether local structure is tubular or non-tubular. The structure classifier is defined as:

$$\xi = \mu_1/(\alpha + \mu_2) - \mu_2/(\alpha + \mu_3), \quad (6)$$

where $\alpha = 0.001$ and $(\mu_1 > \mu_2 > \mu_3)$ are eigenvalues of tensor product in equation (2). $\xi \ll 0$ when the structure is tubular, $\xi \approx 0$ when structure is sphere-like (background and noise), and $\xi \gg 0$ when structure is plate-like. The eigenvalues of HDGS diffusion tensor is combination of eigenvalues of EED and CED:

$$\lambda_{\mu_i} = (1 - \varepsilon)\lambda_{c_i} + \varepsilon\lambda_{e_i}, \quad (7)$$

$$\varepsilon = \exp\left(\frac{\mu_2(\lambda_h^2(\xi - |\xi|) - 2\mu_3)}{2\lambda_h^4}\right), \quad (8)$$

where λ_h is contrast parameter. When the local structure is tubular, $\varepsilon \rightarrow 0$ and the diffusion is CED-like, for other structures $\varepsilon \rightarrow 1$, and diffusion is EED-like.

VED: Manniesing (2006) used the multi-scale Hessian filter response to drive the diffusion. The main idea of the multi-scale Hessian filter is that by using eigenvalues of Hessian matrix, which determine local curvatures, we can distinguish tubular structures from other structures in a multi-scale framework. The output response is a combination of the maximum responses at each scale. Let $V \in [0,1]$ is the output of a multi-scale scale vesselness filter. V should be around 1 inside tubular structures and 0 elsewhere. Assume that $|\lambda_1| < |\lambda_2| < |\lambda_3|$ are eigenvalues of Hessian matrix H , corresponding to eigenvectors Q_1, Q_2 and Q_3 , i.e., $H = Q \cdot \Lambda \cdot Q^T$. Then, Q_1 should be the direction of the least curvature (along vessel in case of vessel structure). The diffusion tensor D is defined as $= Q \cdot \Lambda' \cdot Q^T$. Diagonal elements of matrix Λ' can be defined as:

$$\begin{aligned} \lambda'_1 &= 1 + (\omega - 1)V^{1/s}, \\ \lambda'_2 = \lambda'_3 &= 1 + (\varepsilon - 1)V^{1/s}, \end{aligned} \quad (9)$$

where s is a sensitivity parameter which controls the impact of V on λ'_1 ; ω is a parameter larger than 1, which ensures that Q_1 is always the direction of largest diffusion; ε is relative small to allow high isotropic diffusion when $V \approx 0$ (non-vessel structure).

3 Experiments and Evaluations

3.1 Data

We randomly chose 7 training datasets and 7 testing datasets from portal venous phase liver CTAs that were acquired in Erasmus Medical Center, Rotterdam. The datasets have an in-plane pixel size of 0.74 mm x 0.74 mm, 1-1.5 mm slice spacing, 1-2 mm slice thickness, 72-180 axial slices, 512 x 512 pixels per slice. Scanning was performed 60 seconds after the injection of 100 cc intravenous contrast agent with radiation dose of 140 - 320 mAs and control voltage of 80 - 120 kV. The datasets contain portal veins, hepatic veins, tumors, cysts, metastases in different shape and size. Some datasets also contain metal artifacts.

To reduce filtering times, we cropped the datasets into smaller datasets which containing all liver vessels, which resulted in a reduction of around 75 %.

3.2 Evaluation Criteria

We use SNR as a quantitative metric to determine whether the images have improved. We calculate the SNR over a large set of point pairs that were annotated in the vessel and in the background. We follow the following protocol to calculate the SNR:

1. We choose 12 to 15 random axial slices within liver region to ensure that number of vessels in those is sufficient (more than 300 vessels per data)
2. In every slice, a trained observer annotated all vessels, which have diameter from 0.74mm (one pixel) to around 10 mm, by clicking one point in the vessel, we call them object points.
3. For each object point, a corresponding nearby points in the liver parenchyma is manually selected, which is at least 5 pixels away from the vessels. We call these background points.
4. The SNR of each pair is calculated as:

$$SNR = 20 \log_{10} \frac{obj - mean(bgr)}{std(bgr)} \text{ (dB)}, \quad (10)$$

where the standard deviation of the background points is determined in an axial 5x5 ROI around the point.

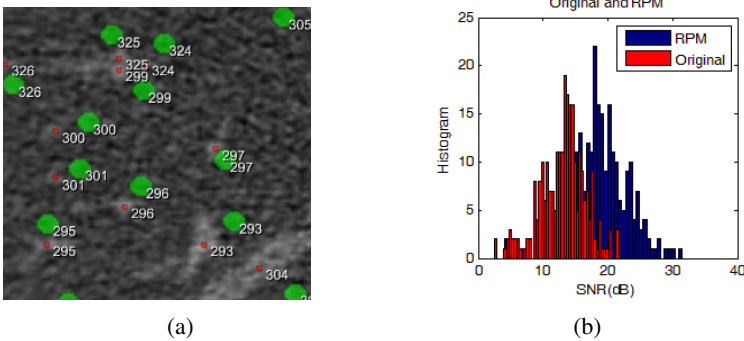


Fig. 1. (a) The vessel markers for SNR calculation: Small red dots are vessel markers, respectively indexed to big green circles on nearby background areas. (b) SNR histograms of an original image (red) and a diffusion image (gray).

From the SNR of all these pairs in one datasets, we determine the mean and standard deviation, which we use to quantitatively evaluate the filter's results and analyze the effect of the filters on large and small vessels.

3.3 Parameter Optimization

The optimal filter parameters are determined in a training stage. For each of the training datasets, we apply the five filters and tune parameters to get the optimal result, based on the mean SNR over all datasets. For the range of values for each parameter we follow the suggestions of Mendrik et al. [10].

The parameter optimization is performed on Linux cluster which contain 80 2.4GHz-64 bits-cores. For RPM, CED, EED, it takes 10 to 20 seconds per iteration,

while for CED and VED, it takes 30 to 60 seconds per iteration. The maximum number of iteration of all filters is set to 50. All parameter values are described in detail in appendix A.

4 Results

4.1 Optimal Parameters

Fig. 2(a) shows the optimal contrast parameter of EED and Fig. 2(b) presents the curve of SNR versus number of iterations for RPM. Fig 2(a) shows the optimization results for the contrast parameter for EED, which demonstrates that the value of that parameter is dataset – independent, whereas the number of iterations [Fig 2(b)] is dataset-dependent: the more noisy data is, the more number of iteration is required.

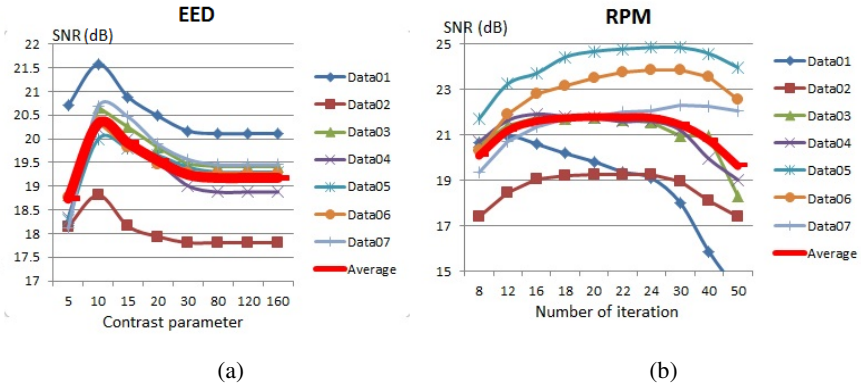


Fig. 2. (a) The average optimal contrast parameter for EED is at 10 (the highest SNR value of the red curve). (b) The average optimal number of iterations for RPM is around 22.

The optimal parameters resulting from the training step are listed in Table 1.

Table 1. The optimal parameters of the diffusion filters

Filter	Abbr	Optimal Parameters
Regular Perona- Malik	RPM	$\tau = 0.0625; \eta = 22; \lambda = 10; \sigma = 1$
Edge-Enhancing Diffusion	EED	$\tau = 0.0625; \eta = 40; \lambda_c = 10; \sigma = 1$
Coherence-Enhancing Diffusion	CED	$\tau = 0.0625; \eta = 50; \lambda_c = 5; \alpha = 0.001; \sigma = 1; \rho = 1$
Hybrid Diffusion filter with Continuous Switch	HDCS	$\tau = 0.0625; \eta = 40; \lambda_c = 5; \lambda_e = 10; \lambda_h = 10; \alpha = 0.001; \sigma = 1; \rho = 1$
Vessel Enhancing Diffusion	VED	$\tau = 0.0625; \eta = 32; \omega = 25.0; s = 1; \varepsilon = 0.01; \sigma_{\min} = 1; \sigma_{\max} = 3; v = 5; \alpha = 0.5; \beta = 0.5; \gamma = 120$

4.2 Filter Comparison

We applied the filters on the 7 test datasets with the optimal parameters from Table 1. Ranking of each diffusion filter are based on average SNR over the test datasets. The

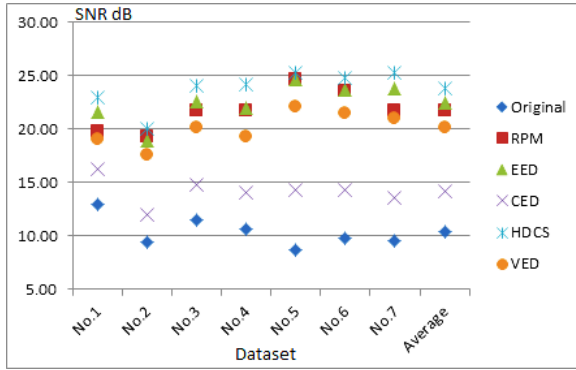


Fig. 3. Ranking on test dataset

results of the overall performance are summarized in Fig. 3. According to the results, HDCS performed better than the other filters.

From Fig. 3, we can see that all filtered results have better SNR than original ones. This also can be seen as SNR histogram in Fig. 1. Fig. 1(b) is an example of SNR histograms of an original image (red) and a diffusion image (gray). The diffusion image’s histogram is to the right of the original’s one, that means the filter has improved SNR in general.

To visualize the effect of diffusion filters in terms of SNR, we calculated the changes in SNR between every pair of points of the filtered images and those of the original ones. The delta SNR is divided into 3 parts: the red part is smaller than zero, that means at those locations, the diffusion results in a worse SNR.; the yellow part is from zero to mean of SNR, which means that at those locations, the diffusion filter is able to improve SNR; the green points are those where the SNR is larger than the mean of the SNR, which means that there is a large improvement. The results for one dataset are shown in Fig. 4.

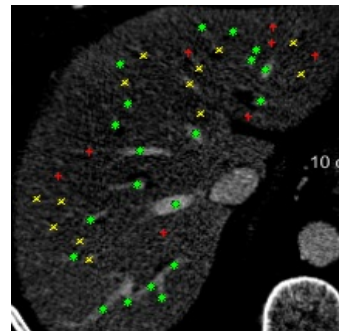
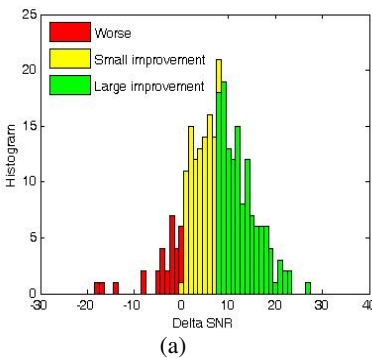


Fig. 4. (a) Histogram of Delta SNR and (b) vessel marker render, respectively. The yellow part is remarked by yellow x characters; The green part is noted by green asterisks; The red part is represented by red + characters.

5 Discussion

We have optimized and applied five diffusion filters for increasing the SNR in liver vessel CTA images. In the parameter optimization stage, we choose a time step which is quite small, $\tau = 0.0625$, to ensure stability of diffusion scheme. In the experiments, we can see that, when increasing the number of iteration, SNR increases as noise is suppressed. However, when the number of iterations becomes too large, both vessel structure and noise are blurred, which results in an SNR decrease. Around the optimal number of iteration, the SNR curve is quite flat [Fig. 2(b)].

To visually compare the results, outputs of each filter and the unfiltered image are shown in Fig. 5 with the same window-level, we can see that the original image has good contrast but much noise. All of the filters, to some extent, blurred low-contrast smaller-than-1 mm vessels. The reason may be that, in optimization step, at Gaussian scale $\sigma = 1$, large vessels have more SNR improvement than the SNR reduction in the small vessels [Fig. 4 (b)]. Diffusion filters blur noise, improve high contrast vessels but also blur low-contrast small vessel.

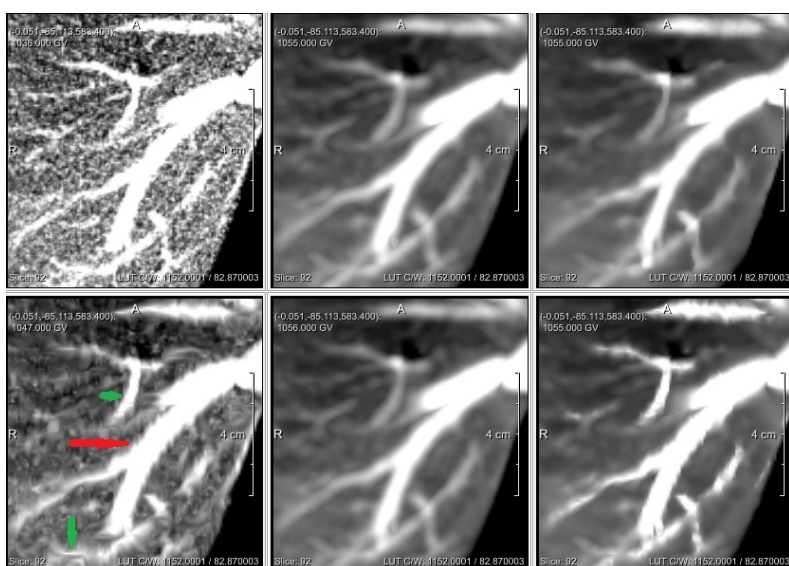


Fig. 5. Maximum intensity projection (7 slices) of the diffusion outputs. Order of top row from left to right: Original, RPM, EED. Order of bottom row from left to right: CED, VED, HDCS.

Also in Fig. 5, EED, VED, and RPM reduce contrast between vessels and background. This effect results from blurring in all directions if gradient or curvature is not high enough to prevent blurring. CED and HDCS not only keep better contrast between vessel and background, but also retain better the structure of small vessels (short-blue green). However, in this study, CED leads to irregular borders of large vessels (long-red arrow). The main reason is that noise at the edges makes the CED tensor consider this noise as small vessels. For HDCS, this effect is not as much as CED because at boundary, EED has some impacts. This effect doesn't influence SNR in general because we only take evaluate the intensity at the center of vessel. In this

study, we used coherent factor $\kappa = (\mu_1 - \mu_2)^2 + (\mu_2 - \mu_3)^2 + (\mu_3 - \mu_1)^2$ for CED as suggested in the original paper by Weickert [1999]. This results in rough effect at boundary of big vessel. In the HDCS paper [10], Mendrik introduced a new coherent factor $\kappa = (\mu_2/(\alpha + \mu_3))^4$. This causes the boundaries of large vessels to be smoother compared to the CED version in this study. In further study, we could use that factor to have better adapted for 3D images.

In test stage, CED shows the worse ranking [Fig 3]. The main reason of this ranking is that CED mainly blurs inside the vessel while noise in flat areas is not suppressed much. In contrast, HDCS not only performs well in side vessel as CED does, but also uses the diffusion property of EED, blurring when gradient is small, in flat areas. This quantitative result is in agreement with the conclusion of the qualitative evaluation by Mendrik et al.

This study just shows optimal parameters for global SNR. However at optimal SNR, all the diffusions expose problem with very small and low contrast vessels. Depending on specific clinical application, for instant in RFA (when detection of small vessels is relevant), we can use this study to recognize very small and low contrast vessels, and then optimize setting again.

6 Conclusion

We presented a quantitative evaluation of five diffusion filters, RPM, EED, CED, HDCS and VED on 3D CTA images of the liver. We optimized the relevant parameters of each filter on a training set of seven CTAs. Based on an evaluation on an independent set of seven datasets and using SNR as criterion, we conclude that HDCS filter performs the best over the other filters.

References

1. Lesage, D., Angelini, E.D., Bloch, I., Funka-Lea, G.: A Review of 3D Vessel Lumen Segmentation Techniques: Models, Features and Extraction Schemes. *Med. Image Anal.* 13, 819–845 (2009)
2. Freiman, M., Joskowicz, L., Sosna, J.: A Variational Method for Vessels Segmentation: Algorithm and Application to Liver Vessels Visualization. In: *Proc. of SPIE*, vol. 7261, p. 72610H (2009)
3. Drechsler, K., Laura, C.O.: Comparison of Vesselness Functions for Multiscale Analysis of the Liver Vasculature. In: *10th IEEE International Conference on Information Technology and Applications in Biomedicine*, pp. 1–5. IEEE Press, New York (2010)
4. Erdt, M., Raspe, M., Suehling, M.: Automatic 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)
5. Lehmann, K.S., Ritz, J.P., Valdeig, S., Schenk, A., Holmer, C., Peitgen, H.O., Buhr, H.J., Frericks, B.B.: Portal Vein Segmentation of a 3D-Planning System for Liver Surgery - In Vivo Evaluation in a Porcine Model. *Ann. Surg. Oncol.* 15, 1899–1907 (2008)
6. Selle, D., Preim, B., Schenk, A., Peitgen, H.O.: Analysis of Vasculature for Liver Surgical Planning. *IEEE Trans. Med. Imaging* 21, 1344–1357 (2002)

7. Alhonnoro, T., Pollari, M., Lilja, M., Flanagan, R., Kainz, B., Muehl, J., Mayrhauser, U., Portugaller, H., Stiegler, P., Tscheliessnigg, K.: Vessel Segmentation for Ablation Treatment Planning and Simulation. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 45–52. Springer, Heidelberg (2010)
8. Manniesing, R., Viergever, M.A., Niessen, W.J.: Vessel Enhancing Diffusion: A Scale Space Representation of Vessel Structures. *Med. Image Anal.* 10, 815–825 (2006)
9. Weickert, J.A.: Coherence-Enhancing Diffusion Filtering. *Int. J. Comput. Vis.* 31, 111–127 (1999)
10. Mendrik, A.M., Vonken, E.J., Rutten, A., Viergever, M.A., van Ginneken, B.: Noise Reduction in Computed Tomography Scans Using 3-D Anisotropic Hybrid Diffusion with Continuous Switch. *IEEE Trans. Med. Imaging* 28, 1585–1594 (2009)

Appendix A

Parameters design:

Table 2. The parameters in optimization stage: σ (Gaussian scale); C (contrast constant); λ , λ_c , λ_e (contrast parameters to RPM, CED and EED); τ (time step); η : (number of iteration).

Filter	Abbr	Parameter optimization values
Regular Perona- Malik	RPM	τ : 0.0625; σ : 0.5, 1; τ : 0.0625 η : 4, 8, 12, 16, 18, 20, 22, 24, 30 40, 50 ; λ : 5, 10, 12, 14, 16, 18, 20, 25, 30, 80, 150;
Edge-Enhancing Diffusion	EED	σ : 0.5, 1; C: 3.31488; τ : 0.0625 λ_c : 5, 10, 30, 80, 120, 160 η : 4, 8, 12, 16, 20, 24, 30, 35 ,40 45, 50
Coherence-Enhancing Diffusion	CED	σ : 0.5, 1; ρ : 0.5, 1; τ : 0.0625; α : 0.001 λ_c : 5,10, 30, 80, 120, 160 η : 4, 8, 12, 16, 22, 26, 30, 35, 40, 45, 50
Hybrid Diffusion filter with Continuous Switch	HDCS	σ : 0.5, 1; ρ : 0.5, 1; α : 0.001 λ_h : 5, 10, 15, 20, 30 λ_c : 5; λ_e : 10 η : 4, 8, 12, 16, 20, 25, 30, 40, 50
Vessel Enhancing Diffusion	VED	σ : 1 – 3 (5 scales); τ : 0.0625 α : 0.5; β : 0.5; γ : 10, 40, 80, 120, 160, 280 η : 4, 8, 12, 16, 20, 25, 32, 40, 50 ω : 8, 25, 32 s : 0.5, 1, 2, 5 ε : 0.01; c : 10^6

[†]Contrast parameters in HDCS can be used form EED and CED.

3D Shape Analysis for Liver-Gallbladder Anatomical Structure Retrieval

Weimin Huang¹, Wei Xiong¹, Jiayin Zhou¹, Jing Zhang⁵, Tao Yang¹, Jiang Liu¹, Yi Su², Calvin Lim², Chee Kong Chui³, and Stephen Chang⁴

¹ Institute for Infocomm Research, Agency for Science, Technology and Research,
1 Fusionopolis Way, #21-01 Connexis South Tower, Singapore 138632, Singapore
{wmhuang, wxiong, jzhou, tyang, jliu}@i2r.a-star.edu.sg

² Institute of High Performance Computing, Agency for Science, Technology and Research,
1 Fusionopolis Way, #16-16 Connexis North Tower, Singapore 138632, Singapore
{suyi, limcw}@ihpc.a-star.edu.sg

³ Department of Mechanical Engineering, Faculty of Engineering,
National University of Singapore, Centre for Biomedical Materials Application
and Technology, 9 Engineering Drive 1, Singapore 117576, Singapore
mpecck@nus.edu.sg

⁴ Department of Surgery, National University Hospital,
1E Kent Ridge Road, Singapore 119228, Singapore
cfscky@nus.edu.sg

⁵ Department of Medical Information Engineering,
School of Electrical Engineering and Information, Sichuan University, P.R. China
dzzj2001@gmail.com

Abstract. Anatomical structure is important for medical education and disease diagnosis. In the application of surgical simulation, different anatomical structures can be retrieved to create variety of surgical scenarios for training, while similar structures can also be retrieved to assist disease diagnosis. This paper presents an approach to liver-gallbladder anatomical structure retrieval with 3D shape comparison, where the direct shape comparison based on dense shape registration is applied to liver shape due to its shape complexity, and feature based comparison is applied to gallbladder shape with a semantic shape decomposition using the saliency area based on multi-scale curvatures and concavity. After the registration of liver models, the geometric structure of the gallbladder and liver can be combined for joint comparison. With the 3D models constructed from a set of liver-gallbladder CT data, experiments are conducted for joint liver-gallbladder retrieval. Encouraging result shows that it can reveal important topology based on similarity and variance of 3D shapes and has a similar performance compared to that of manual retrieval by human operators.

Keywords: Anatomical structure, shape analysis, shape comparison, surgical simulation.

1 Introduction

Anatomical relationships of liver and gallbladder (GB), of which some are rare such as the left GB and intrahepatic GB, may pose challenges to the surgeons. The GB

shape information, such as extra large size of a GB or folds on a GB, is not only useful for surgery and training, but may also reveal some potential pathological risks. In laparoscopic surgery and simulation of cholecystectomy, anatomical shape and structure of the liver and GB, including the position, size and wall thickness, are some of the important factors that affect the difficulty of such surgeries. Most of the current surgery simulation systems mainly aim at the training on fundamental laparoscopic skills (FLS) with generic liver, GB and other organ models. Such a simulation thus lacks the facility to supply different scenarios with different challenges for training. However it is possible to construct new models from patient CT data and store the models in a model library for future training as well as patient specific surgery planning and preoperative practice.

The training with new models can assist the surgeon trainees by exposing them with new scenarios and different challenges instead of a single generic model. With increasing number of models, it becomes difficult for a trainee to select suitable training cases, to cater for his/her specific training purposes. Thus a reliable approach to compare and retrieve the relevant (similar or dissimilar) anatomical structure is highly needed.

1.1 Related Works

GB diseases have varying symptoms, which sometimes can be shown in the shape change of a GB [1, 2]. Although the shape anomaly may not always be related to disease, it may cause the stasis that could lead to stone formation and inflammation [3]. Measurement for wall thickness from CT images has been studied by Prasad *et al* [4]. Noticeably, GBs can also vary from person to person in terms of size, shape and location, which pose the challenges for new surgeons.

In cholecystectomy, liver is the main surrounding organ of GB. Thus liver shape is to be modeled in the simulation and the anatomical relationship specifies the surgery scenario. In the literature, liver shape model has been successfully applied for atlas construction [5] and segmentation [6]. However there is no study on the retrieval of anatomical structure of liver-GB. The statistical liver shape model and liver shape retrieval has been successfully applied for segmentation [7]. However there is no systematic study on the retrieval of liver-gallbladder. Work in [8] presented physics modeling of GB, which did not address GB shape analysis. In [9] the researchers described an approach to a 3D shape decomposition, which is improved using a more robust saliency method in this paper.

In 3D model retrieval, works include shape description, modeling and similarity measurement [10]. Generally a 3D shape can be represented by tetrahedral or polyhedral mesh. Focusing on surface model, there are three main categories of 3D shape analysis. One is based on the mesh patch segmentation from the salient points on the surface [11], and the second is based on skeletonization of 3D shapes or logical parts decomposition [12]. Besides, there are some works proposed to compare the shape based on the registered models directly using the distance of the corresponding points [13, 14]. To segment an object into meaningful parts, minima rule and part salience [15] have been proposed. Banegas *et al* presented a decomposition of volume data

[16], and then ellipsoids are fitted and deformed to reconstruct the object from the hierarchical decomposition.

1.2 Contributions

This paper presents an approach to the application in joint liver-GB shape comparison and retrieval, where direct shape registration and comparison based on Coherent Point Drift (CPD) [13] is adopted to compute liver shape similarity and semantic features are proposed for GB shape comparison. Using CPD is due to its speed and nonlinearity for non-rigid shape registration. Combining the distance measurement of the liver-GB anatomical structures, users can retrieve the similar or dissimilar data sets for surgery training or pathological comparisons. Noted that the liver-gallbladder pairs are essentially in a continuous shape space, as no categories are defined for joint liver-gallbladder anatomical structure, the conventional precision-recall measurement can not be adopted for the retrieval performance evaluation. Instead, the Spearman's score for rank correlation [20] is used to compare the order of the retrieval data sets with the result from human operators, which can be the trainee surgeons or technicians who are selecting similar or dissimilar anatomical structures for training.

2 Liver-Gallbladder Model Comparison

The preprocessing of CT data to reconstruct liver and GB model is based on our previous work for CT image segmentation and mesh construction and optimization [17], [18]. This paper focuses on liver-GB shape analysis and retrieval, where we compare liver shape based on direct shape distance due to the complexity of liver shape, while model GBs with meaningful semantic features.

2.1 Similarity of Liver Shapes

The registration method adopted here is based on CPD where point set registration is formulated as a probability density estimation problem. The moving set is modeled by a Gaussian mixture model (GMM) with unknown centroids. It forces the GMM centroids to move coherently as a group by reparameterizing the centroid locations explicitly. The form of transformation is non-rigid. Since we choose to use a fast Gaussian transformation [13] the registration is fast (about 10 seconds for 2 sets of 1000 points/vertices on surface mesh). Given the registration of liver shapes L_1 and L_2 , we compute the mean Hausdorff distance of the two shapes, modified (multiplied) by dot product of the surface normals of the two correspondences.

Fig. 1 shows the superimposed liver shapes, from similar shape to dissimilar ones. Only rigid transformation is used for better visualization so the original shapes can be compared visually.

To quantitatively evaluate the performance of the proposed shape comparison, we use the Spearman's score that computes the rank correlation [20] on two of the shape similarity ranking orders. We get the Spearman's score on liver retrieval between

human operators and the method. As the liver shapes are highly complex, the shape similarity rank is quite subjective from person to person especially for dissimilar shapes comparison. Nevertheless, for those similar shapes, the rank should be consistent.

Here, the Spearman's score ρ between the two human operators is only 0.40. Analyzing the result we found that for the dissimilar liver sets the retrieval results were quite subjective, which caused the low correlation among the sets. While for those liver shapes that have many similar samples, the correlation rank is higher, the score $\rho=0.53$ if four sets that have less similarity with each other are not used in the comparison. The correlation score by using the CPD shape registration with the two human operators are increased to 0.54 and 0.47 respectively.

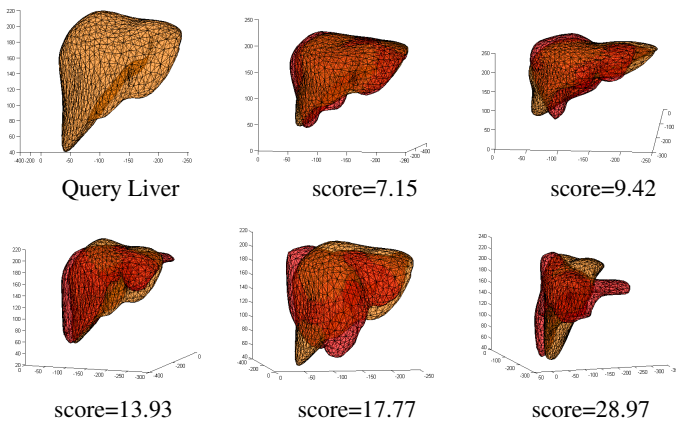


Fig. 1. Liver registration and score using modified mean Hausdorff distance measurement

2.2 Gallbladder Shape Modeling

GB shape can be intuitively and semantically represented by neck, size, orientation, and folds. To detect a GB fold, we use saliency region to segment a GB. We have developed a method for the decomposition [9]. However it has problems in ring-like saliency region caused by shape protrusion or large saliency region caused by a flattened fold. Here we present a new approach to find the proper cutting plane. With a surface mesh, a saliency region R_s at scale s is defined as a set of connected mesh vertices V , of which all the Gaussian curvatures K_s are negative due to the opposite values of the two principal curvatures.

$$R_s = \{v \in V, K_s(v) < 0\}. \quad (1)$$

To capture the global shape change, we detect all local minimum in R_s of Gaussian curvatures at largest scale N , $V_m = \{v_m \in V, K_N(v_m) < K_N(v_{r1}), v_{r1}$ is the neighboring vertices at ring 1 of $v_m\}$. As K_N is negative, we have

$$v_p = \arg \min_{v_m} K_n(v_m) C_m(v_m). \tag{2}$$

The concavity C_m measures the minimum distance from the mesh vertex v_m to the mesh convex hull. It is a good measurement of the global shape concavity at the vertex. To consider the saliency region around v_p (2), set the curvature threshold as half of the local minima, $\xi_s=0.5K_s(v_p)$, $R_s = \{v \in V, K_s(v) < \xi_s\}$. A multiscale saliency region R_p is then defined as

$$R_p = \cup R_s, \text{ and } \cap R_s \neq \emptyset. \tag{3}$$

Let the saliency region be Q containing v_p . The center of the saliency region is the mean,

$$v_c = \frac{1}{N_Q} \sum_{v_i \in Q} v_i, \tag{4}$$

where Q has N_Q vertices. Define the main norm n_c for the saliency region as the surface norm nearest to v_c . The principal (curvature) directions of the surface at a vertex v_i in Q can be computed [19]. Shown in Fig. 2 are the curvature directions, with the blue lines being the minimum principal curvature directions p^n , and yellow lines being those for maximum principal direction p^m . Red dots are the saliency points forming the saliency region (on left image). Note that we are interested in a cutting plane aligned with n_c . Illustrated in Fig. 2 (central image), let the cutting plane have a normal n_p , v_c is a point on the plane, then $n_c \times n_p = 0$. Removing the n_c component from p^n , we have,

$$q^n = p^n - (p^n \cdot n_c) n_c. \tag{5}$$

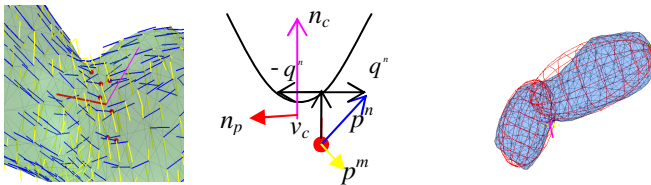


Fig. 2. Cutting plane using principal curvature direction. Left: partial surface at a saliency region, Middle: the illustration of the curvatures, Right: the ellipsoid fitting for a GB.

For a vertex, both p^n and $-p^n$ are possible depending on the surface property. Randomly select a q_0 we adjust the other direction q_i by aligning them with q_0 ,

$$q_i^n = \begin{cases} q_i^n, & \text{if } q_i^n \cdot q_0^n \geq 0 \\ -q_i^n, & \text{if } q_i^n \cdot q_0^n < 0 \end{cases} \quad (6)$$

The cutting plane normal is then computed as

$$n_p = \frac{1}{N_e} \sum_i q_i^n \quad (7)$$

With the cutting plane, the mesh vertices of a GB are separated into two sets of data points. One ellipsoid is applied to one set of points by the ellipsoid fitting.

The GB neck is further identified by a position related to the liver center. Let a GB mesh vertices be $G=\{g_i\}$, the liver mesh vertices be $L=\{l_j\}$, centered at l_c . The major principal axis O_g of G can be obtained by PCA. Mapping all $\{g_i\}$ to O_g , so $d_i = \{(g_i - g_c) \cdot O_g\}$. The GB neck N_e is obtained as one of the two extremes $\{g_{e1} = \arg \max(d_i), g_{e2} = \arg \min(d_i)\}$ that is close to the liver center l_c . Another extreme is taken as the fundus. Here, given N_e , a GB is simplified by the connection of the major axis of the decomposed ellipsoids, (E_1, E_2) , pointing from the GB neck to the GB fundus. Fig. 3 illustrates such a GB topology, which is the semantic feature of the GB, written as $G = \{N_e, O_g, S_g, (E_1, E_2)\}$, where S_g is the size of the GB.

Depending on fold detection, a GB may have only one ellipsoid if no saliency region detected. For GBs A and B both with two ellipsoids (E^A_1, E^A_2) , and (E^B_1, E^B_2) , the bend difference (angle and size) is used to measure the similarity. By shifting N^A_e to N^B_e , and rotating the ellipsoids (E^A_1, E^A_2) around $E^A_1 \times E^B_1$ so that E^A_1 is aligned to E^B_1 , the distance of the folds is,

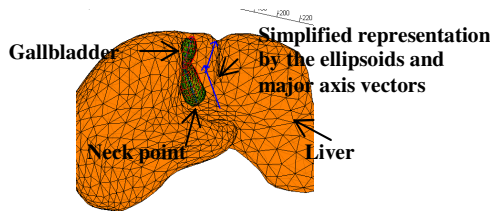


Fig. 3. Gallbladder representation

$$D_b = 1 - w_{AB} \frac{|E^A_2 \cdot E^B_2|}{\|E^A_2\| \times \|E^B_2\|} \quad (8)$$

The weight w_{AB} measures the difference of relative length l between the two folds,

$$w_{AB} = \min(\|I^A\|, \|I^B\|) / \max(\|I^A\|, \|I^B\|), \quad (9)$$

$$I_1^G = \|E_1^G\| / (\|E_1^G\| + \|E_2^G\|). \tag{10}$$

If a GB G with only one ellipsoid E_1^G , letting $E_2^G = E_1^G$, (8) still works well as a distance measurement. It can be shown that if the folds are the same, $D_b=0$.

2.3 Liver-Gallbladder Distance Measurement

In this study, liver volume is normalized so that it will not affect the comparison. It can always be added easily later whenever it is necessary. Mahalanobis distance is adopted to compute the joint liver-GB distance between patients, p_1 and p_2 . Let D^L be the distance between liver shapes, and $D_{i=1,\dots,4}^G = (D_{N_e}, D_{o_g}, D_{s_g}, D_b)$ be the distance between the semantic features of the GBs. To balance the shape influence of liver and GB, the joint similarity is rewritten as

$$S(p_1, p_2) = 1 - \frac{1}{2} \left(\frac{|D^L|}{\sigma^L} + \frac{1}{4} \sum_i \frac{|D_i^G|}{\sigma_i^G} \right). \tag{11}$$

The σ in (11) is the corresponding standard deviation.

3 Experiments and Result Analysis

Experiments have been conducted on a small data set including 19 liver-GB CT scans. Liver volume is normalized so that it will not affect the comparison. By visual examination of the results in Fig. 4, we can see that a high similarity score (11) does reflect the anatomical similarity, both in shapes and structures. The gallbladders are superimposed at the bottom-right for a better view. Query-1 has obtained very similar anatomical structures on top. Query 2 has only one similar instance, but due to the gallbladders' shapes are not similar, the overall score is low.

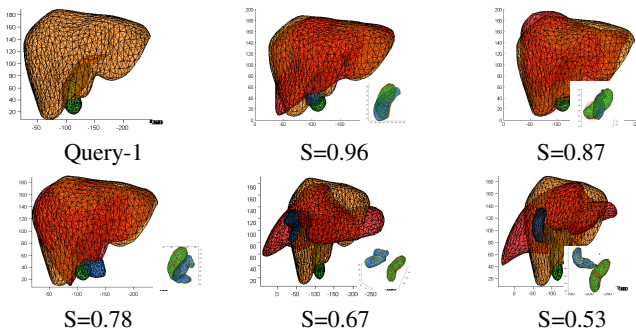


Fig. 4. Liver-GB retrieval with similarity scores

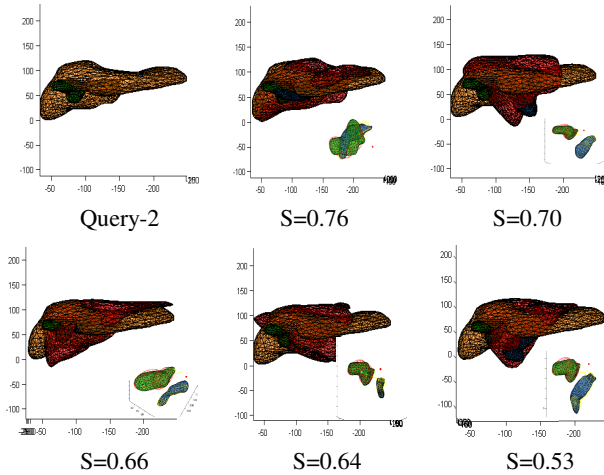


Fig. 4. (continued)

Spearman's rank correlation [20] is also used to measure the performance. The higher the Spearman's score ρ is, the more similar of the two ranks will be. Two users are asked to sort the data based on the visual similarity. The mean score ρ between the two users is 0.55, over all 19 queries. Compared with the results from the two users, the proposed approach got the mean scores $\rho=0.50$ and 0.40 respectively. The main difference is caused by some dissimilar datasets, which are difficult even for human operators to decide the similarity rank among them. If we take out the dissimilar sets (4 sets, irregular liver and gallbladder shapes), the score between the 2 operators is 0.68 and the scores go up to 0.64 and 0.54 between our method and the two human operators. We also tested the shape retrieval method in [21] using multi-view LightField Descriptor. Including all data for testing, the correlation score with human operators are 0.34 and 0.35. One of the reasons is that the LightField is good for similar objects, but it may not suitable for dissimilar objects. In GB comparison, the size and orientation are the factors to be considered, however which will be normalized by LightField.

4 Conclusion

This paper presents the approaches to model and compare 3D anatomical structures of liver-gallbladder pairs. Through jointly comparing the shape and structure features, the liver-gallbladder anatomical retrieval is developed. Although the data set collected and tested is small, it includes variant structures and shape changes like fundus fold, big fold in the middle of gallbladder, normal liver shapes, and abnormal liver shapes (some are due to lesions) with gallbladders in different types and positions. The

preliminary test shows that the approach successfully retrieved the data according to the similarities or dissimilarity. The quantitative results show that our method is highly correlated with human's performance.

The structure retrieval proposed in the paper is limited to either similar or dissimilar cases for the training purpose based on liver shape, GB shape and the relationship between liver and GB. The shape registration using CPD or other generic registration does not consider the semantic similarity of livers, such as the similarity between liver left/right lobes or the GB fossa shapes, which requires specific region identification on the lobes, segments and fossa. For GB comparison, only the biggest fold is detected and used for GB decomposition. For GBs with more complicated folds, the method may not be enough to characterize the shape variation. Currently the detected fold can be a real fold or a small 'fold' caused by the extension from GB to cystic duct. In the comparison (8), we do not distinguish the types of folds, but the relative length and position of the folds are used to measure the difference between two GBs.

Future works include the benchmark on more datasets, improvement using relevance feedback, and incorporation of other structures for data retrieval.

Acknowledgements. This project is funded by the Agency for Science, Technology and Research (A*STAR), Singapore under BEP Grant 1021480009.

References

1. Meilstrup, J.W., Hopper, K.D., Thieme, A.: Imaging of Gallbladder Variants. *Am. J. Roentgenol.* 157(6), 1205–1208 (1991)
2. Bodzioch, S.: Automated Detecting Symptoms of Selected Gallbladder Illness Based on A Static Ultrasound Images Analysis. *Bio-Alg. and Med-Sys.* 2, 35–44 (2006)
3. Kaiser, E.: Congenital and Acquired Changes in Gallbladder Form. *Am. J. Dig. Dis.* 6(7), 938–953 (1961)
4. Prasad, M.N., Brown, M.S., Ni, C., Margolis, D.J., Douek, M., Raman, S., Lu, D., Goldin, J.: Three-Dimensional Mapping of Gallbladder Wall Thickness on Computed Tomography Using Laplace's Equation. *Acad. Radiol.* 15, 1075–1081 (2008)
5. Xiong, W., Ong, S.H., Tian, Q., Xu, G., Zhou, J., Liu, J., Venkatash, S.K.: Construction of a Linear Unbiased Diffeomorphic Probabilistic Liver Atlas from CT Images. In: *IEEE International Conference on Image Processing*, pp. 1773–1776. IEEE Press, New York (2009)
6. Okada, T., Shimada, R., Sato, Y., Hori, M., Yokota, K., Nakamoto, M., Chen, Y.-W., Nakamura, H., Tamura, S.: Automated Segmentation of the Liver from 3D CT Images Using Probabilistic Atlas and Multi-level Statistical Shape Model. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 86–93. Springer, Heidelberg (2007)
7. Chi, Y., Cashman, P., Bello, F., Kitney, R.I.: An Automatic Liver Segmentation Initialization Information Retrieval Strategy for a Content-Based Image Retrieval System Followed by a New Liver Volume Segmentation Method for CT and MRI Image Datasets. In: *MICCAI 2007 Workshop on Content-Based Image Retrieval for Biomedical Image Archives* (2007)

8. Zhang, J., Huang, W., Zhou, J., Yang, T., Liu, J., Su, Y., Chui, C.K., Chang, S.: Gallbladder Modeling and Simulation in Laparoscopic Cholecystectomy. In: IEEE International Conference on Industrial Electronics and Applications, pp. 128–131. IEEE Press, New York (2011)
9. Huang, W., Zhou, J., Liu, J., Zhang, J., Yang, T., Su, Y., Law, G.H., Chui, C.K., Chang, S.: 3D Shape Decomposition and Comparison for Gallbladder Modeling. In: Proc. of SPIE Medical Imaging, vol. 7964, p. 79642K (2011)
10. Tangelder, J.W.H., Velkamp, R.C.: A Survey of Content Based 3D Shape Retrieval Methods. *Multimed. Tools Appl.* 39, 441–471 (2008)
11. Yamauchi, H., Gumhold, S., Zayer, R., Seidel, H.-P.: Mesh Segmentation Driven by Gaussian Curvature. *Visual Comput.* 2, 659–668 (2005)
12. Au, O.K.-C., Tai, C.-L., Chu, H.-K., Cohen-Or, D., Lee, T.-Y.: Skeleton Extraction by Mesh Contraction. *J. ACM Trans. on Graphics* 27(3), no. 44 (2008)
13. Myronenko, A., Song, X.: Point Set Registration: Coherent Point Drift. *IEEE Trans. Pattern Anal. Mach. Intell.* 32(12), 2262–2275 (2010)
14. Besl, P.J., McKay, N.D.: A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* 14(2), 239–256 (1992)
15. Shapira, L., Shamir, A., Cohen-Or, D.: Consistent Mesh Partitioning and Skeletonisation Using the Shape Diameter Function. *Visual Comput.* 24, 249–259 (2008)
16. Banegas, F., Jaeger, M., Michelucci, D., Roelens, M.: The Ellipsoidal Skeleton in Medical Applications. In: ACM Symposium on Solid Modeling and Applications, pp. 30–38 (2001)
17. Zhou, J., Xiong, W., Ding, F., Huang, W., Qi, T., Wang, Z., Oo, T., Venkatesh, S.K.: Liver Workbench: A Tool Suite for Liver and Liver Tumor Segmentation and Modeling. In: Loménie, N., Racoceanu, D., Gouaillard, A. (eds.) *Advances in Bio-imaging*. AISC, vol. 120, pp. 193–207. Springer, Heidelberg (2012)
18. Su, Y., Chua, K.S., Chong, C.S.: Mesh Processing Using Virtual Geometry. *WSEAS Transactions on Computers* 5(4), 696–704 (2006)
19. Cohen-Steiner, D., Morvan, J.-M.: Restricted Delaunay Triangulations and Normal Cycle. In: ACM Symposium on Computational Geometry, pp. 237–246 (2003)
20. Vogel, J., Schiele, B.: Semantic Modeling of Natural Scenes for Content-Based Image Retrieval. *Int. J. Comput. Vis.* 72(2), 133–157 (2007)
21. Chen, D.-Y., Tian, X.-P., Shen, Y.-T., Ouhyoung, M.: On Visual Similarity Based 3D Model Retrieval. In: *Computer Graphics Forum (EUROGRAPHICS 2003)*, vol. 22(3) (2003)

Tumor Sensitive Matching Flow: An Approach for Ovarian Cancer Metastasis Detection and Segmentation

Jianfei Liu¹, Shijun Wang¹, Marius G. Linguraru², and Ronald M. Summers¹

¹ Imaging Biomarkers and Computer-Aided Diagnosis Laboratory,
Radiology and Imaging Sciences, National Institutes of Health Clinical Center
Bethesda, MD 20892, USA
jianfei.liu@nih.gov

² Sheikh Zayed Institute for Pediatric Surgical Innovation,
Children's National Medical Center,
Washington DC 20010, USA

Abstract. Accurately detecting and segmenting ovarian cancer metastases can have potentially great clinical impact on diagnosis and treatment. The routine machine learning strategies to locate ovarian tumors work poorly because the tumors spread randomly to the entire abdomen. We propose a tumor sensitive matching flow (TSMF) to identify metastasis-caused shape variance between patient organs and atlas. TSMF juxtaposes the role of feature computation/classification, and TSMF vectors highlight tumor regions while dampening all other areas. Therefore, metastases can be accurately located by choosing areas with large TSMF vectors, and segmented by exploiting the level set algorithm on these regions. The proposed algorithm was validated on contrast-enhanced CT data from 11 patients with 26 metastases. 84.6% of metastases were successfully detected, and false positive per patient was 1.2. The volume overlap of the segmented metastases was $63 \pm 5.6\%$, the Dice coefficient was $77 \pm 4.2\%$, and the average surface distance was $3.9 \pm 0.95\text{mm}$.

Keywords: Ovarian cancer metastases, computer-aided detection, tumor sensitive matching flow.

1 Introduction

Detecting and segmenting ovarian cancer metastases enhance the prognosis and treatment of women with ovarian cancer because 75% of them have tumors that have already spread at the time of diagnosis [9]. The metastasis detection manifests many challenges, including 1) variability in shapes and locations among individuals, 2) indistinctive intensity profile in comparison with surrounding tissues, 3) abnormal shapes of human organs compressed by metastases. Unpredictable locations of metastases hinder existing detection algorithms [12,13,4]

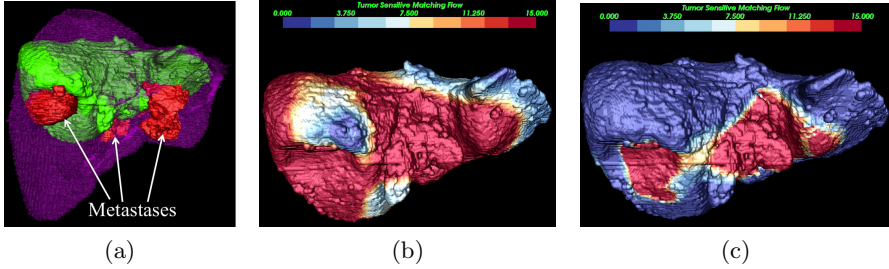


Fig. 1. (a) Significant shape variance between the registered atlas (purple) and the patient’s liver (green); (b) optical flow methods can track the shape variance, which is mapped to the patient’s liver. Blue to red represents small to large displacements; (c) tumor sensitive matching flow highlights the displacements at tumor regions while dampening all other deformed areas.

to classify ovarian tumors despite the fact that salient tumor classifiers can be trained on annotated datasets. Manual annotation and classifier training are usually time-consuming. Moreover, most detection algorithms [13, 6] focus on finding lesions or tumors inside organs. Our purpose is instead to locate *exterior ovarian cancer metastases attached to organs*. Therefore, metastasis detection without training would be desirable.

Metastases can be alternatively located by measuring local shape variance between patient data and atlas because they often push organs to deform. Image registration serves this purpose. For instance, free-form deformation registration [10] exploits the spline model to track non-rigid motion. Fig. 1a illustrates the shape variance between the registered atlas liver (purple) and the patient’s liver (green). The difference is partly due to attached metastases (red), but mainly caused by the variability among individuals. Therefore, shape comparison solely based on image registration is unreliable. Optical flow methods [2, 7] can compute relative image displacements between registered and patient livers, as shown in Fig. 1b. The amounts of image displacements are color mapped to the patient’s liver. Red areas contain large displacements, which match actual shape variance. However, the individual variability dominates the image displacements while metastases take minor effects.

In this paper, we study the problem of eliminating shape variance caused by individual variability while keeping the variance due to metastases, so as to identify them. We propose a tumor sensitive matching flow (TSMF) to integrate local tumor classifier into optical flow computation. Tumor-like regions are emphasized during flow computation, while all other areas are suppressed. Fig. 1c illustrates the results of TSMF, where red regions correspond to the locations of metastases. Finally, we can place level set seeds in these areas to segment them.

2 Methodology

The flow of our method for detecting and segment metastasis is described in Fig. 2. It consists of three major steps: shape descriptor construction, tumor sensitive matching flow computation, and metastasis segmentation.

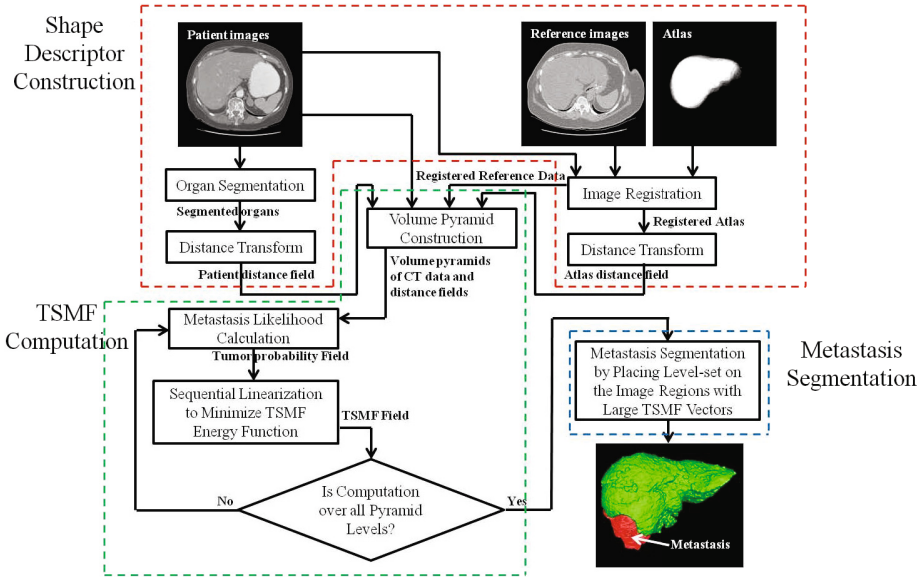


Fig. 2. Tumor sensitive matching flow method for detecting and segmenting metastasis

2.1 Shape Descriptor Construction

Shape is the primary feature for radiologists to detect metastasis. The purpose of this step is to build shape descriptors for both patient images and reference CT images (scanned from healthy persons), depicted in top left and center images of Fig. 2.

Liver and spleen are segmented using the method in Linguraru [5] from patient images because ovarian cancer metastases frequently attach to them. Distance transform [8] is then performed on the segmented organs to build the distance field, which is the shape descriptor for the patient data. Similar process is performed on one reference CT dataset. The Reference CT images are first registered with the patient images [10]. The registration parameters are then used to transform a probabilistic atlas, shown in top right image of Fig. 2, to the patient coordinate. The registered atlas thus covers the possible spatial ranges

that the healthy organ would have. Another distance field is calculated from the registered atlas and used as the shape descriptor for the reference images. Therefore, we obtain two pairs of datasets: patient and registered reference images, and patient and atlas distance fields.

2.2 Tumor Sensitive Matching Flow (TSMF) Computation

The TSMF computation is the key to accurately identify metastases by comparing two pairs of datasets from the previous step. Let $I_p(x, y, z)$ and $I_a(x, y, z)$ be the patient images and reference images, and $D_p(x, y, z)$ and $D_a(x, y, z)$ be their corresponding distance fields, with $\vec{u} = (u_x, u_y, u_z)$ be the TSMF vector at point (x, y, z) . Similar to optical flow computation [2,7], the TSMF computation can be formulated as a global energy function within a minimization framework

$$\begin{aligned}
 E(\vec{u}) = & \iint_{(x,y,z) \in \mathbb{R}^3} \underbrace{\Psi((I_a(x + u_x, y + u_y, z + u_z) - I_p(x, y, z))^2)}_{\text{Intensity Constancy}} \\
 & + \beta G(x, y, z) \underbrace{\Psi((\nabla I_a(x + u_x, y + u_y, z + u_z) - \nabla I_p(x, y, z))^2)}_{\text{Gradient Constancy}} \\
 & + \gamma G(x, y, z) \underbrace{\Psi((D_a(x + u_x, y + u_y, z + u_z) - D_p(x, y, z))^2)}_{\text{Distance Constancy}} \\
 & + \alpha \underbrace{\Psi(|\nabla u_x|^2 + |\nabla u_y|^2 + |\nabla u_z|^2)}_{\text{Flow Smoothness}} dx dy dz,
 \end{aligned} \tag{1}$$

where $\Psi(x^2) = \sqrt{x^2 + \epsilon^2}$, $\epsilon = 0.001$ is a modified $L1$ norm and allows the computation to handle non-Gaussian deviations of the matching criterion. α , β , and γ are constants to balance different components. $G(x, y, z)$ is a metastasis-likelihood equation for estimating the probability of the metastasis existence at point (x, y, z) . The larger the value of $G(x, y, z)$, the more influence the distance and gradient constancy terms will conduct in the local flow computation. Therefore, flow vectors are magnified at the locations where metastasis is more likely to exist.

Next, we clarify the definition of $G(x, y, z)$. In Fig. 2.2, we notice that the intensity values of the metastasis are slightly lower than the liver and its region is approximately homogeneous. A Gaussian kernel is used to model the intensity distribution of metastases, and we experimentally determine that $\mu_m = 1060\text{HU}$ and $\sigma_m = 20\text{HU}$ are the average and standard deviation from one representative dataset. The metastasis also generates the local concavity of the liver, and it can be measured by $S(x, y, z) = D_a(x, y, z) - D_p(x, y, z)$. Moreover, it stays at the exterior of the liver, and thus $G(x, y, z)$ should be a piecewise function that highlights metastases at organ's boundaries. Let Ω and $\bar{\Omega}$ be the segmented

organs and non-organ regions, respectively. $\partial\Omega$ is the organ boundary. $G(x, y, z)$ is defined as

$$G(x, y, z) = \begin{cases} 0.01 & \text{for } (x, y, z) \in \overline{\Omega} \\ S(x, y, z)/\exp(\frac{\sigma}{\sigma_m})(1 + (\frac{\mu - \mu_m}{\sigma_m})^2) & \text{for } (x, y, z) \in \partial\Omega \\ 0.01 \times (D_{max} - D_p(x, y, z)) & \text{for } (x, y, z) \in \Omega. \end{cases} \quad (2)$$

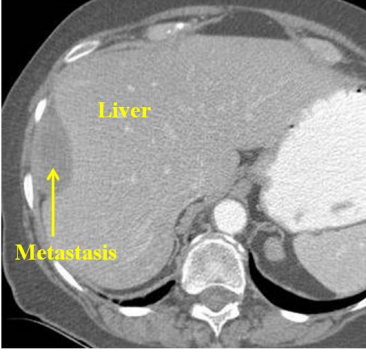


Fig. 3. The analysis of metastasis properties to design a metastasis-likelihood function

Here, D_{max} is the largest distance value at the patient distance field, and μ and σ are mean and deviation of the intensity values of the non-organ pixels adjacent to the current boundary point in the patient images. Equation (2) indicates that the likelihood of the metastasis existence remains a small value in the non-organ regions and gradually decreases towards the organ. The likelihood significantly increases if the local boundary has large shape change as well as the intensity values remains homogeneous within the metastasis intensity level. Therefore, Equation (2) is sensitive to the metastases attaching to organs.

Minimization. Equation (1) is non-trivial to be minimized because it is a highly non-linear and non-convex equation. In order to ease the description, we define the following [7]:

$$\begin{aligned} \Delta I &= I_a(x + u_x, y + u_y, z + u_z) - I_p(x, y, z) \\ \Delta D &= D_a(x + u_x, y + u_y, z + u_z) - D_p(x, y, z) \\ \Delta(\partial_x I) &= \partial_x I_a(x + u_x, y + u_y, z + u_z) - \partial_x I_p(x, y, z) \\ \Delta(\partial_y I) &= \partial_y I_a(x + u_x, y + u_y, z + u_z) - \partial_y I_p(x, y, z) \\ \Delta(\partial_z I) &= \partial_z I_a(x + u_x, y + u_y, z + u_z) - \partial_z I_p(x, y, z) \end{aligned} \quad (3)$$

In terms of the calculus of variations, the Euler-Lagrange equation regarding to x component is expressed as

$$\begin{aligned} &\Psi'((\Delta I)^2)\partial_x I_a \Delta I + \beta G(x, y, z)\Psi'((\Delta(\partial_x I))^2 + (\Delta(\partial_y I))^2 + (\Delta(\partial_z I))^2) \\ &(\partial_{xx} I_a \Delta(\partial_x I) + \partial_{xy} I_a \Delta(\partial_y I) + \partial_{xz} I_a \Delta(\partial_z I)) + \gamma G(x, y, z)\Psi'((\Delta D)^2)\partial_x D_a \Delta D \\ &- \alpha \operatorname{div}(\Psi'(|\nabla u_x|^2 + |\nabla u_y|^2 + |\nabla u_z|^2)\nabla u_x) = 0. \end{aligned} \quad (4)$$

The equations of y and z components can be similarly derived. However, Equation (4) is still nonlinear in its argument \vec{u} .

Multi-scale analysis is an efficient approach to handle non-convexity of Equation (4) as the solution in the coarse scale can better approximate the global minimum. Volume pyramids are constructed to simulate scale space on patient

and reference images as well as their distance fields. Sampling rate 0.75 is used to ensure the smooth transition between different scales.

Sequential linearization [2,7] is another numerical strategy to remove non-linearity in Equation (4). It is represented as two nested fixed-point iterations. Assuming k be pyramid level and l be the outer iteration index,

$$(\Delta I)^{k,l+1} = (\Delta I)^{k,l} + (\partial_x I_a)^{k,l} du_x^{k,l} + (\partial_y I_a)^{k,l} du_y^{k,l} + (\partial_z I_a)^{k,l} du_z^{k,l}, \quad (5)$$

where $\vec{u}^{k,l+1} = (u_x^{k,l} + du_x^{k,l}, u_y^{k,l} + du_y^{k,l}, u_z^{k,l} + du_z^{k,l})$. Accordingly, non-linearity at ΔI is iteratively removed, and the same strategy can be performed on other abbreviations in Equation (3). Let the inner iteration index be m , the purpose of the inner iteration is to manipulate all $\Psi'(*)$ operators only relying on $d\vec{u}^{k,l,m}$ when $d\vec{u}^{k,l,m+1}$ is being estimated. Therefore, Equation (4) is finally converted into a linear equation after two nested iterations.

Successive over-relaxation method [14] is employed to minimize the massive linear system over the entire volume. After two nested iterations exceed predefined values at the current pyramid level, the solutions are used as the initialization for the next pyramid level through bilinear interpolation. TSMF field is generated after the computation is accomplished at the finest pyramid level. Fig. 1c shows the final results.

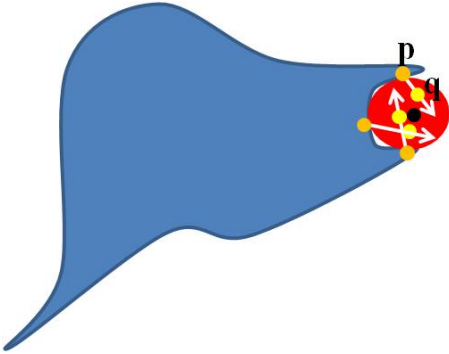


Fig. 4. The process of seed point determination, where the metastasis is represented as a red sphere and the liver is shown in blue. A transitional point q is determined by moving a surface point p with half length of the flow vector, indicated as white arrows. The seed point is determined by choosing the center points of a set of transitional points satisfying the intensity requirement.

2.3 Metastasis Segmentation

Level set method based on fast marching [11] is employed to segment metastases due to its accuracy. The key to extracting metastases successfully is the determination of a set of seed points.

TSMF field makes the task of seed determination tractable. Potential metastasis regions are first extracted by selecting organ surfaces with the length of flow vectors larger

than 15mm. Connecting graphs are then built on the selected surfaces to determine the number of connected regions, in other words, the number of potential metastases. If the number of vertices of a connected region exceeds 100 (approximately $100mm^2$ area on the organ surface), it is evenly split.

Fig. 4 illustrates the process of seed point determination within a connected region. Let $\mathbf{p} = (x, y, z)$ be a surface point, represented as a green point. We first compute the transitional point $\mathbf{q} = (\hat{x}, \hat{y}, \hat{z}) = (x + \frac{u_x}{2}, y + \frac{u_y}{2}, z + \frac{u_z}{2})$ in yellow. A set of transitional points can be obtained, and its center point is $\bar{\mathbf{q}} = \sum_i \mathbf{q}_i$, indicated as a black point. If $\mu_m - \sigma_m < I(\bar{\mathbf{q}}) < \mu_m + \sigma_m$, $\bar{\mathbf{q}}$ is chosen as a seed point. Otherwise, search its adjacent points and select neighbored points that fulfill the intensity range. If none of them satisfy the condition, this region is rejected.

The determined seed points are then imported into the fast marching approach and metastases are finally segmented.

3 Experimental Results

TSMF algorithm was tested on 11 abdominal contrast-enhanced CT datasets generated by Siemens 64-detector CT scanner. Slice thicknesses was 1mm. Each dataset has at least one ovarian cancer metastasis. Retrospective analysis of these images was inspected by our Institutional Reviewer Board. 26 metastases in the selected datasets were annotated by an experienced radiologist and used as the ground-truth. Their size (the maximum diameter) range is 4.0-49.9mm. 22 metastases were attached to the liver, and the remaining 4 were touched to the spleen.

It takes 20 minutes to process one patient. Fig. 5 illustrates the results from four patients, corresponding to four rows. Ground-truth metastases are illustrated in the right column, TSMF fields are given in the center column, and our segmentation results are shown in the right. The first patient in Fig. 5 has a metastasis in the right side of the liver. The TSMF field accurately tracks the shape change caused by this metastasis. As a result, the metastasis is identified and segmented in the right column. There is one false positive on the gallbladder. The intensity profile and shape of the gallbladder as computed by the TSMF are similar to those of the metastasis. The similar result was observed in the second patient. In the third patient, the spleen is also attached to a metastasis and TSMF can still locate and segment it correctly. The fourth patient is a challenging case because most organs in the left abdomen were removed. Metastasis (A) located at the left abdomen is attached to the liver only at one slice. Because the shape change is minor, TSMF misses it. The same issue happens to the metastasis (B) because of its small size. However, the remaining metastases are successfully detected by TSMF.

The detailed validation of detection and segmentation on 11 patients were presented in table 1. Sensitivity (Sen.) and false positive per patient (FP/Patient) are used to evaluate detection results. Six metrics used in liver segmentation [5] are employed to evaluate the metastasis segmentation. They are volume overlap (VO), Dice coefficient (DC), relative absolute volume difference (RA), average symmetric absolute surface distance (AS), symmetric RMS surface distance (SR), and maximum symmetric absolute surface distance (MS).

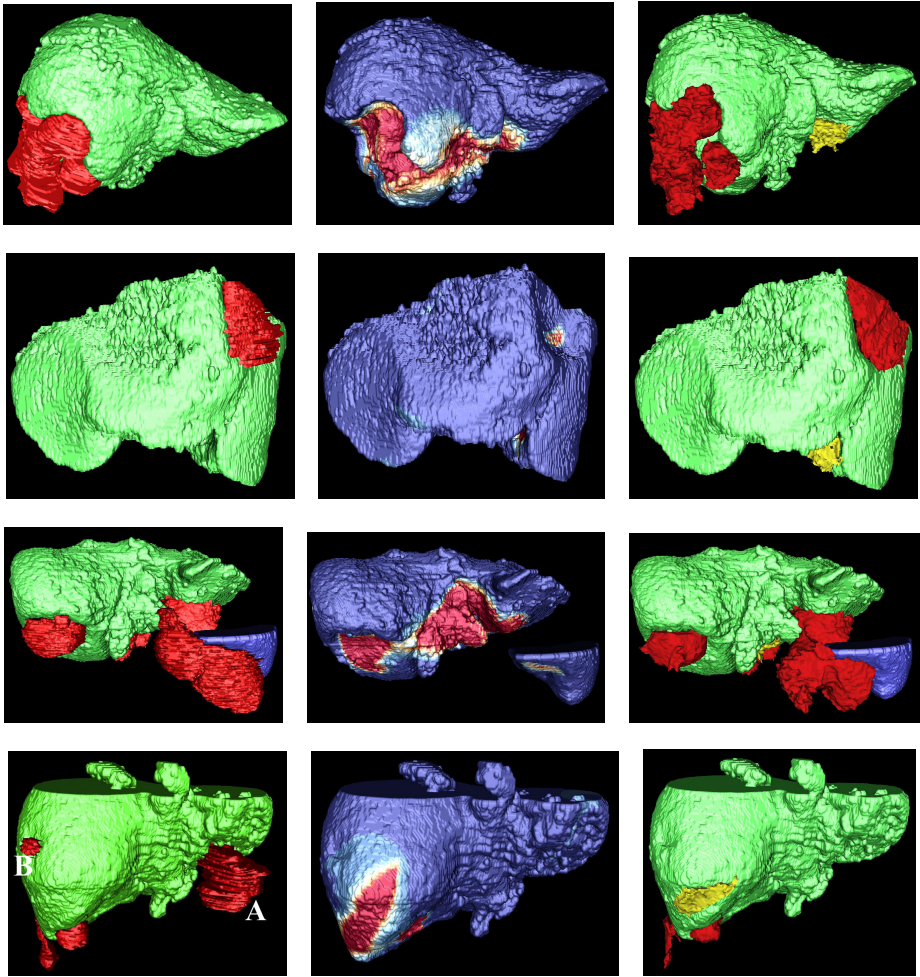


Fig. 5. The comparison between ground-truth (left column) and segmented metastases (right column). TSMF results are also illustrated in the center column, where TSMF vectors are color mapped to the organ surfaces (blue to red represents the increment of flow vectors). Each row corresponds to a patient. True metastases are shown in red and false positives in yellow. False positives tend to be located near the gallbladder because its intensity and shape are similar to metastases. Livers and spleens were automatically segmented. Metastasis A is missed because of incomplete liver segmentation and B is due to its small size (5.6mm).

4 Conclusion and Future Work

We have proposed a tumor sensitive matching flow (TSMF) algorithm to detect and segment ovarian cancer metastases randomly distributed in the abdomen

Table 1. Validation results of metastasis detection and segmentation on 11 patients

Detection		Segmentation					
Sen.(%)	FP/Patient	VO(%)	DC(%)	RA(%)	AS(mm)	SR (mm)	MS (mm)
84.6	1.2	63±5.6	77 ± 4.2	27 ± 15	3.9 ± 0.95	6.5 ± 2.6	15 ± 2.7

from contrast-enhanced CT data. TSMF provides an efficient means to measure shape variance caused by metastases between patient images and atlas data while suppressing all other deformations. Therefore, metastases can be accurately located and segmented according to the TSMF field. The method can successfully detect 84.6% of metastases on data from 11 patients with an average surface distance of 3.88mm.

However, the TSMF presented 13 false positives over 11 patients. Seven of them were located on the gallbladder because it has a similar intensity distribution and shape to the metastases. Currently, we are introducing the gallbladder atlas to the metastasis detection, so as to reduce false positives. Moreover, three of four true positives were missed due to small shape changes between metastases and organs. We are developing a more accurate metastasis-likelihood function based on the information from local image structures, such as texture and shape index, to enhance the sensitivity to the metastases. In addition, segmented metastases sometimes include the surrounding tissues because the boundaries between them are undistinguished. Segmentation strategies based on other image information beyond intensities is being studied to prevent over-segmentation. Last but not least, we are collecting more datasets to test our detection algorithm. Not only are patients with metastases chosen, but also the datasets without metastases are also considered to evaluate the robustness of our detection algorithm.

Acknowledgements. This research was supported by the Intramural Research Program of the NIH Clinical Center. The authors thank Dr. Elise Kohn for helpful comments.

References

1. Bilello, M., Gokturk, S.B.: Automatic detection and classification of hypodense hepatic lesions on contrast-enhanced venous-phase ct. *Medical Physics* 31, 2584–2593 (2004)
2. Brox, T., Bruhn, A., Papenberg, N., Weickert, J.: High Accuracy Optical Flow Estimation Based on a Theory for Warping. In: Pajdla, T., Matas, J.(G.) (eds.) *ECCV 2004, Part IV*. LNCS, vol. 3024, pp. 25–36. Springer, Heidelberg (2004)
3. Gletsos, M., Mougiakakou, S.G., Matsopoulos, G.K., Nikita, K.S., Nikita, A.S., Kelekis, D.: A computer-aided diagnostic system to characterize ct focal liver lesions: design and optimization of a neural network classifier. *IEEE Transactions on Information Technology in Biomedicine* 7, 153–162 (2003)

4. Janowczyk, A., Chandran, S., Singh, R., Sasaroli, D., Coukos, G., Feldman, M.D., Madabhushi, A.: Hierarchical Normalized Cuts: Unsupervised Segmentation of Vascular Biomarkers from Ovarian Cancer Tissue Microarrays. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part I. LNCS, vol. 5761, pp. 230–238. Springer, Heidelberg (2009)
5. Linguraru, M.G., Sandberg, J.K., Li, Z., Shah, F., Summers, R.M.: Automated segmentation and quantification of liver and spleen from ct images using normalized probabilistic atlases and enhancement estimation. *Medical Physics* 37, 771–783 (2010)
6. Linguraru, M.G., Richbourg, W.J., Watt, J.M., Pamulapati, V., Summers, R.M.: Liver and Tumor Segmentation and Analysis from CT of Diseased Patients via a Generic Affine Invariant Shape Parameterization and Graph Cuts. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) Abdominal Imaging 2011. LNCS, vol. 7029, pp. 198–206. Springer, Heidelberg (2012)
7. Liu, J., Subramanian, K.R., Yoo, T.S.: Temporal volume flow: an approach to tracking failure recovery. In: Proc. of SPIE Medical Imaging, lake Buena Vista, Florida, USA (2011)
8. Maurer, J.C., Qi, R., Raghavan, V.: A linear time algorithm for computing exact euclidean of distance transform of binary images in arbitrary dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25, 265–270 (2003)
9. Memarzadeh, S., Berek, J.: Advances in the management of epithelial ovarian cancer. *The Journal of Reproductive Medicine* 46, 621–629 (2001)
10. 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 Transactions on Medical Imaging* 18, 712–721 (1999)
11. Sethian, J.A.: Level set methods: evolving interfaces in computation geometry, fluid mechanics, computer vision, and materials Science, 1st edn. Cambridge University Press (1999)
12. Smeets, D., Loeckx, D., Stijnen, B., Dobbelaer, B.D., Vandermeulen, D., Suetens, P.: Semi-automatic level set segmentation of liver tumors combining a spiral scanning techniques with supervised fuzzy pixel classification. *Medical Image Analysis* 14, 13–20 (2010)
13. Summers, R.M.: Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. *Gastroenterology* 129, 1832–1844 (2005)
14. Young, D.: Iterative Solution of Large Linear Systems (Computer Science and Applied Mathematics), 1st edn. Academic Press (1971)

Fast Segmentation of Abdominal Wall: Application to Sliding Effect Removal for Non-rigid Registration

Wenwu Zhu^{1,2}, Stephane Nicolau¹, Luc Soler¹, Alexandre Hostettler¹,
Jacques Marescaux¹, and Yves Rémond²

¹ IRCAD-France, Virtual-surg, place de l'Hôpital 1, 67091 Strasbourg Cedex, France

² Institut de Mécanique des Fluides et des Solides, rue Boussingault 2,
67000 Strasbourg, France
cool.wenwu@gmail.com

Abstract. The non-rigid registration of abdominal images is still a big challenge due to the breathing motion. Indeed, the sliding between the abdominal wall and the abdominal viscera makes the local deformation field discontinuous; it means that the classical registration approach, which assumes a smooth global deformation field cannot provide accurate and clinical-required results. Other new approaches intend to add in regularization a term to allow discontinuous deformation field near sliding boundary, however, the performance of such approaches needs to be further evaluated. We propose a new approach to perform abdominal image registration including a priori knowledge of the sliding area. Our strategy is to firstly delineate the abdominal wall in source and target images and create new images containing viscera only. Then a state-of-the-art non-rigid registration algorithm is adopted for the registration of the viscera region. In this paper, we firstly show why and how a quick interactive delineation of the full abdominal wall (AW) can be performed using B-spline interpolation. Secondly, we evaluate our registration approach on arterial and venous phase CT images. The results of our approach are compared to the one obtained using the same algorithm with the same parameters on the original data (without segmentation). The registration errors (mean \pm SD) with our approach are: liver (1.94 ± 2.76 mm), left kidney (0.38 ± 0.66 mm), right kidney (0.42 ± 0.82 mm), spleen (4.15 ± 3.68 mm), which is much better than the registration result without segmentation: liver (6.48 ± 10.00 mm), left kidney (3.14 ± 3.39 mm), right kidney (2.79 ± 3.12 mm), spleen (17.45 ± 12.39 mm). The results clearly demonstrate our approach is a promising method to remove the sliding motion effect on the non-rigid registration of abdominal images.

Keywords: Sliding motion, abdominal image registration, image segmentation.

1 Introduction

Image registration plays a significant role for many clinical applications. Though there are lots of registration approaches and algorithms to handle several kinds of situation, abdominal organs registration is still a challenging task since there is not only deformation of the abdominal organs, but there is also their sliding against the abdominal wall (AW) during the breathing motion [1].

Non-rigid image registration of medical image is usually an ill-posed problem: it may give accurate but non-realistic results. One classical approach for registering abdominal images is to adopt a similarity measure metric to get optimization parameters of the transformation model. For example, a B-spline transformation model combined with mutual information similarity metric has been used for breast MR image registration [2], liver motion modeling [3], and liver registration [4]. The main limitation of this approach is that it assumes that the global deformation field is smooth, so it cannot represent the local sliding between moving structures during the breathing cycle.

Thus, to make registration results suitable for clinical applications, regularization terms have been proposed to take the deformation field discontinuities into account. They are based on spatial smoothing of a non-parametric dense deformation field. It includes the use of the anisotropic diffusive regularization for representing the discontinuities at sliding interface [5, 6], and the combination of the demon algorithms with local affine transformation to obtain a smooth diffeomorphic dense deformation field [7-9]. However, the performance of these approaches has not been clearly quantitatively evaluated, and thus need further validation.

After discussion with medical and anatomy experts, it can be summarized that the main sliding motion occurs between the abdominal wall and abdominal viscera. In other words, the sliding motion between abdominal organs themselves is very small compared with the whole abdominal viscera sliding. Therefore, it seems relevant to locate the sliding area and remove the AW from original image before carrying out non-rigid image registration, so that it avoids complex work to adapt a regularization step to artificial take boundary sliding effect into account.

In section 2, we firstly show on patient data that due to the low curvature of abdominal wall, ten slices only need to be delineated in order to provide an accurate full segmentation of the abdominal wall. In section 3, we evaluate on one patient data the benefit of our approach for non-rigid registration of abdominal area.

2 Fast Interactive Segmentation of the Abdominal Wall

In this section, we firstly describe the interactive tool that we have developed in order to quickly delineate abdominal wall in several axial slices. Secondly, we discuss the number of axial slices that must be interactively delineated to reach an accurate segmentation and show on patient data that 10 slices are enough on average.

2.1 Interactive Segmentation Tool Description and Image Generation

Given an abdominal 3D image, the outline of our segmentation is that we select some axial slices to delineate interactively with control points linked by a 2D B-spline. Then the remaining slices are segmented automatically using a 3D B-spline based interpolation technique. Finally, users have to make some small modifications for a more accurate segmentation by adjusting control points of the 2D B-spline curve. The detailed process is as follows.

Given a selected axial slice image, we add control points sequentially along the boundary of the considered region, here the abdominal viscera. Once all control points are added on the boundary, a 2D B-spline based fitting technique is adopted to form a curve linking all adjacent points (cf. Fig.1 (a)). The B-splines being updated in real-time, the place of control points can be adjusted in order to improve the created curve accuracy. Once this step is finished, we go to the next selected slice image. The closest created curve is automatically copied on the current slice. In order to facilitate matching between points from slice i and slice $i + 1$, the control point number of the curve is preserved, therefore, we do not have to append control points again (cf. Fig.1 (b)). Some small modifications on location of control points are sufficient to update the curve which delineates the abdominal viscera. After delineation of at least 4 slices, the boundary of abdominal viscera in the rest of all axial slices can be estimated automatically with the use of 3D B-spline based interpolation technique. Simultaneously, a 3D mesh can also be created using the Marching Cubes algorithm (cf. Fig.1 (c)) and allows to check the segmentation accuracy. Finally, the image to be registered is processed by replacing the AW voxel value by 0 (cf. Fig. (d)).

However, the number of selected slices (NSS) needs to be discussed, because an abdominal volume data normally contains more than 200 slices. If we choose too many slices for interactive segmentation, for example 80% of all slices, it would take an expert more than 5 hours to do the segmentation. If too few slices are chosen, the segmentation result might not be satisfying, we have to find the good compromise.

2.2 Evaluation of the Minimum NSS to Delineate the Abdominal Wall

The length of acquired abdominal-thorax volume data along z axis is usually about 50 cm. Using 8 patient CT data (resolution of $512 \times 512 \times 292$ with voxel size of $0.961 \times 0.961 \times 1.8$ mm), our medical staff perform segmentation with our segmentation tool for each of them using 50, 20, 10 and 5 slices. An expert had defined that the segmentation with 50 slices (1 slice/cm) could be considered as ground truth due to the low curvature of the abdominal wall. Then, a dense 3D mesh M_{50} (resp. S_{20} , S_{10} , S_5) were created for each patient from the segmentation with 50 slices (resp. 20, 10, 5). The surface models S_{20} , S_{10} and S_5 can be compared to the M_{50} and the difference between them can be measured by computing the distance from the vertex of the surface model to the closest point belonging to M_{50} . The number of control points in each slice is the same for the segmentation with different NSS.

Fig. 3 shows the distance histogram between meshes S_{20} , S_{10} and S_5 with M_{50} on 2 patient data. We can see that the distance distribution of points is almost the same for S_{20} and S_{10} (peak around 0.6 mm). But for NSS equals to 5, there are many points which distance are larger than 0.8 mm. We also calculate the total average distance and standard deviation between surfaces S_{20} , S_{10} , and S_5 with M_{50} on 8 patient data (see Table 1). It clearly shows that the mean error is reduced from 1.27 mm (NSS = 5) to 0.84 mm (NSS = 10) which corresponding to the voxel size and thus to the ground truth accuracy. However, there is only slight improvement in accuracy for the NSS increased from 10 to 20.

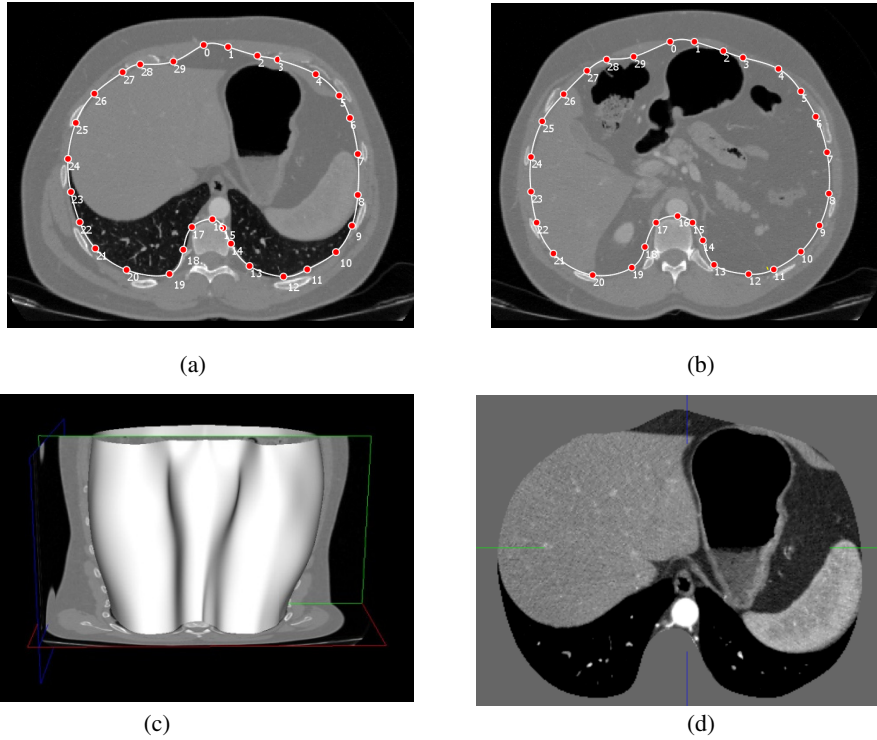


Fig. 1. (a) A curve created interactively with 2D B-spline linking all control points in one slice. (b) A new curve was created which control points are copied from a closest curve (c) A 3D mesh corresponding to the segmentation in 5 axial slices (NSS=5). (d) One axial slice of the new image to be registered.

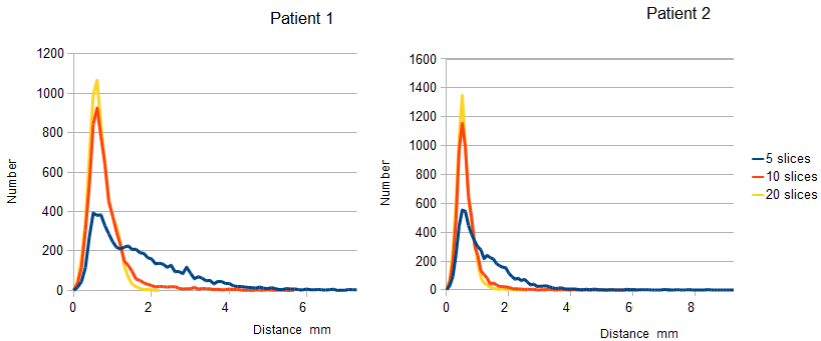


Fig. 2. Distance histogram between S_{20} , S_{10} , S_5 with M_{50} . The x axis represents the distance (mm) and the y axis represents the number of points (vertex) in corresponding distance.

From these statistic charts and tables, we can conclude that the NSS for interactive segmentation can be chosen below 20, but above 10. This evaluation also shows that the abdominal wall can finally be fully segmented in a very reasonable time compared with clinical workflow using semi interactive tool with similar accuracy than ground truth. Particularly, 15 slices are segmented in less than 10 minutes with our software.

Table 1. Average distance and standard deviation between surfaces S_{20} , S_{10} , S_5 with ground truth M_{50}

	5 slices	10 slices	20 slices
Average distance (mm)	1.27	0.84	0.71
SD (mm)	0.99	0.52	0.36

3 Evaluation of Abdominal Wall Removal on Deformable Registration

In this section, we apply state-of-the-art non-rigid algorithm to register an arterial phase abdominal CT image with a venous phase CT image of the same patient. In a first step, we use original image I_{arterial} and I_{venous} , and in a second step we use I'_{arterial} and I'_{venous} which abdominal wall are removed with our proposed method (NSS=15). Finally, we compare the registration result accuracy using several insight measurements (cf. Section 3.4).

3.1 Image Data

The original data we use is two abdominal CT images, namely arterial and venous-phase CT of the same patient. The size is $512 \times 512 \times 292$ with spacing of $0.961 \times 0.961 \times 1.8 \text{ mm}^3$. Then, these two images were segmented by the proposed approach to get the result images without abdominal wall.

3.2 Non-rigid Image Registration

The goal in contrast-enhanced CT abdominal images registration is to find a spatial transformation that correctly maps points in the source image to the corresponding ones in the target image. One important issue that needs to be considered is that not only shape of organs is deformed because of breathing, they also move in cranial-caudal direction. Obviously, using rigid or affine transformation only is not sufficient to accurately describe this transformation. Therefore, we choose a combined transformation T which contains rigid transformation, affine transformation and multi-resolution deformable transformation using free form deformation (FFD) :

$$T(x, y, z) = T_{\text{localDeformable}}(T_{\text{globalAffine}}(T_{\text{globalRigid}}(x, y, z))) \quad (1)$$

The great advantage and capacity of B-spline based FFD compared with thin-plate splines in modelling the 3D deformation field has been proved by [2, 11, 12].

3.3 Choice of Similarity Metric

For the registration of arterial and venous-phase CT images, it is necessary to define a registration metric. Intrinsic value difference between these two data sources make the intensity similarity measure impossible using methods such as sum of squared differences or correlation. We decide to choose the Mutual Information (MI) which is a concept of information theory. MI means how much information one random variable contains about another random variable and does not need a bijection of grey level between source and target images. This was introduced into the medical image registration area by Collignon [14] and Viola [15]. In our paper, we adopt the class `itk::MattesMutualInformationImageToImageMetric` to implement the similarity measurement. More details can be obtained via accessing the ITK website [16].

3.4 Evaluation of the Registration

To evaluate the registration results on original images and on images modified with our approach, several organs (liver, kidney, spleen) have been segmented manually by an expert in the target image (arterial- phase CT) and the same organs have been segmented in the images after registration. These segmentation Then a quantitative analysis is performed. Firstly, for two segmented organs masks (V_1 and V_2), we define the Sensitivity and Specificity as following:

$$S_{sensitivity} = \frac{V_1 \cap V_2}{V_1}, \quad S_{specificity} = \frac{V_1 \cap V_2}{V_2}. \quad (2)$$

In the liver case, V_1 represents the liver mask in the original CT image, V_2 represents its mask in the registered image with or without the abdominal wall. Secondly, Table 2 gives the volume of organs (liver, left/right kidney, spleen). Thirdly, we calculate the mean distance between both masks for each organ, and the results are shown in the Table 3. Here, the mean distance is the average of the distance between all points of the boundary of the organ in original image and their closest point that in the boundary of registered one with or without abdominal wall.

Table 3 shows that the mean distance is significantly reduced when using image in which AW has been removed. For example, the mean distance of the liver without registration is 30.90 mm, then, this value is reduced from 6.48 mm to 1.94 mm corresponding to the registration with and without the abdominal wall. The sensitivity and specificity of all organs are also dramatically improved when AW is not in the registered image. Finally, the important improvement of our registration approach can be assessed on the 3D model in Fig. 3. However, one can see that there are still some errors (5 mm) close to the diaphragm. This is mainly due to the grey level of lungs and heart which are not homogeneous.

Table 2. Organ volume computed from the original data and registered image with and without abdominal wall

Organs	V_{original} (mm ³)	$V_{\text{regWithAW}}$ (mm ³)	$V_{\text{regWithoutAW}}$ (mm ³)
Liver	2004.43	2345.50	2116.32
Left kidney	212.32	200.285	218.508
Right kidney	212.832	198.066	215.948
Spleen	322.944	355.892	354.630

Table 3. Quantitative analysis for the registration accuracy of the liver, left kidney, right kidney and spleen

Organs		Sensitivity	Specificity	Mean (mm)
Liver	NoRegistration			30.90
	RegWithAW	92.51%	79.05%	6.48 (± 10.00)
	RegWithoutAW	96.96%	91.83%	1.94 (± 2.76)
Left kidney	NoRegistration			24.76
	RegWithAW	81.44%	86.33%	3.14 (± 3.39)
	RegWithoutAW	86.33%	96.01%	0.38 (± 0.66)
Right kidney	NoRegistration			24.57
	RegWithAW	82.98%	89.17%	2.79 (± 3.12)
	RegWithoutAW	98.01%	96.63%	0.42 (± 0.82)
Spleen	NoRegistration			30.72
	RegWithAW	38.94%	35.33%	17.45 (± 12.39)
	RegWithoutAW	82.72%	75.33%	4.15 (± 3.68)

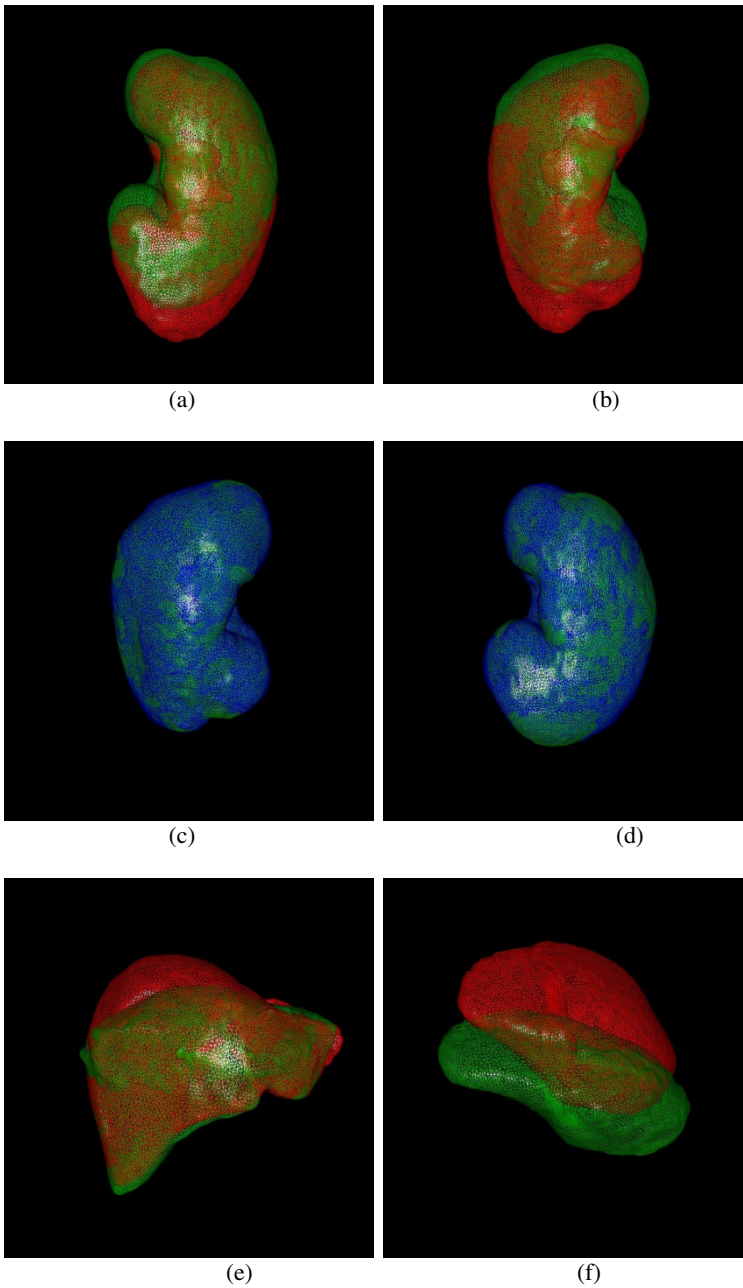


Fig. 3. Visualization of the mesh models of segmented organs in target and registered images. The model colored in green is from target image; the model colored in red is from registered image with abdominal wall; the model colored in blue is from registered image without abdominal wall.

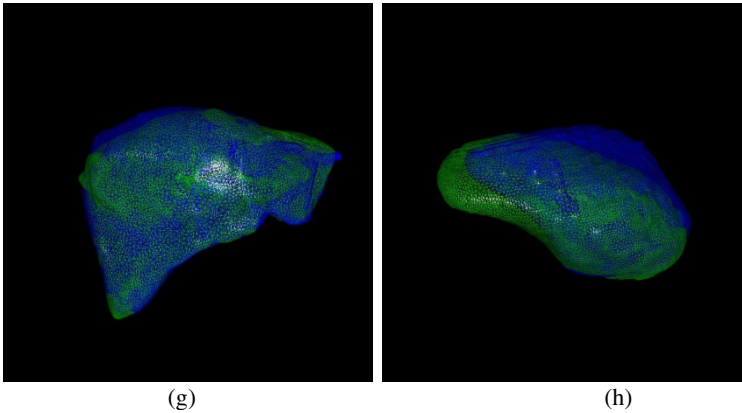


Fig. 3. (continued)

4 Conclusion

In this paper, we have presented a novel and simple approach for the abdominal image registration which is based on a prior and fast delineation of the abdominal wall and a new image generation. It is different from traditional methods which try to build a local regularization field to represent the discontinuity caused by the sliding motion. In fact, our prior delineation can be seen as a priori knowledge related to the main sliding area. We firstly demonstrate that a quick delineation of AW is possible in a very short time using segmentation of a few slices only (typically 10). Secondly, a quantitative evaluation and comparison was performed with the registration of images with AW and the result indicates that our approach provides much higher accuracy. In the future, we will investigate the benefit of segmenting the diaphragm area and we will complete the evaluation on more patient data and from different modalities (typically MRI).

Acknowledgements. This research work is supported by the China Scholarship Council for 3 years PhD study at the University of Strasbourg.

References

1. Kiriyanthan, S., Fundana, K., Cattin, P.C.: Discontinuity Preserving Registration of Abdominal MR Images with Apparent Sliding Organ Motion. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) *Abdominal Imaging 2011*. LNCS, vol. 7029, pp. 231–239. Springer, Heidelberg (2012)
2. Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D.: Non-Rigid Registration Using Free-Form Deformations: Application to Breast MR Images. *IEEE Trans. Med. Imaging* 18(8), 712–721 (1999)

3. Rohlfing, T., Maurer Jr., C.R., O'Dell, W.G., Zhong, J.: Modeling Liver Motion and Deformation during the Respiratory Cycle Using Intensity-Based Free-Form Registration of Gated MR Images. In: *Medical Imaging: Visualization, Displaying, and Image-Guided Procedures*. SPIE, vol. 4319, pp. 337–348 (2001)
4. Lee, W.C., Tublin, M., Chapman, B.: Registration of MR and CT Images of the Liver: Comparison of Voxel Similarity and Surface Based Registration Algorithms. *Comput. Meth. Prog. Bio.* 78, 101–114 (2005)
5. Pace, D.F., Enquobahrie, A., Yang, H., Aylward, S.R., Niethammer, M.: Deformable Image Registration of Sliding Organs Using Anisotropic Diffusive Regularization. In: *8th IEEE International Symposium on Biomedical Imaging*, pp. 407–413. IEEE Press, New York (2011)
6. Pace, D.F., Niethammer, M., Aylward, S.R.: Sliding Geometries in Deformable Image Registration. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) *Abdominal Imaging 2011*. LNCS, vol. 7029, pp. 141–148. Springer, Heidelberg (2012)
7. Freiman, M., Voss, S., Warfield, S.: Demons Registration with Local Affine Adaptive Regularization: Application to Registration of Abdominal Structures. In: *8th IEEE International Symposium on Biomedical Imaging*, pp. 1219–1222. IEEE Press, New York (2011)
8. Cahill, N.D., Noble, J.A., Hawkes, D.J.: A Demons Algorithm for Image Registration with Locally Adaptive Regularization. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part I*. LNCS, vol. 5761, pp. 574–581. Springer, Heidelberg (2009)
9. Freiman, M., Voss, S.D., Warfield, S.K.: Abdominal Images Non-rigid Registration Using Local-Affine Diffeomorphic Demons. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) *Abdominal Imaging 2011*. LNCS, vol. 7029, pp. 116–124. Springer, Heidelberg (2012)
10. Zitova, B., Flusser, J.: Image Registration Methods: A Survey. *Image Vision Comput.* 21(11), 977–1000 (2003)
11. Lee, S., Wolberg, G., Shin, S.Y.: Scattered Data Interpolation with Multilevel B-Splines. *IEEE Trans. Vis. Comput. Graph.* 3, 228–244 (1997)
12. Bardinet, E., Cohen, L.D., Ayache, N.: Tracking and Motion Analysis of the Left Ventricle with Deformable Superquadrics. *Med. Image Anal.* 1(2), 129–149 (1996)
13. Lee, S., Wolberg, G., Chwa, K.Y., Shin, S.Y.: Image Metamorphosis with Scattered Feature Constraints. *IEEE Trans. Vis. Comput. Graph.* 2, 337–354 (1996)
14. Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Seutens, P., Mar, G.: Automated Multimodality Image Registration Using Information Theory. In: *14th International Conference on Information Processing in Medical Imaging, IPMI 1995*, pp. 263–274 (1995)
15. Viola, P., Wells, W.M.: Alignment by Maximization of Mutual Information. *Int. J. Comput. Vis.* 24, 137–154 (1997)
16. Insight Segmentation and Registration Toolkit (ITK), <http://www.itk.org>

A Fully Automated Framework for Renal Cortex Segmentation

Xinjian Chen¹, Hemin Zhao¹, and Jianhua Yao²

¹ School of Electronics and Information Engineering, Soochow University, Suzhou, Jiangsu 215006, China

² Department of Radiology and Imaging Sciences, National Institute of Health Clinical Center, Bldg. 10, Rm 1C224D MSC 1182, Bethesda, Maryland 20892-1182, USA
xjchen@suda.edu.cn

Abstract. The current procedure of renal cortex segmentation is subjective and tedious. This investigation is to develop and validate an automated method to segment renal cortex on contrast-enhanced abdominal CT images. The proposed framework consists of four parts: first, an active appearance model (AAM) is built using a set of training images; second, the AAM is refined by live wire (LW) method to initialize the shape and location of the kidney; third, an iterative graph cut-oriented active appearance model (IGC-OAAM) method is applied to segment the kidney; Finally, the identified kidney contour is used as shape constraints for renal cortex segmentation which is also based on IGC-OAAM. The proposed method was validated on a clinical data set of 27 CT angiography images. The experimental results show that: (1) an overall cortex segmentation accuracy with overlap error $\leq 12.7\%$, volume difference $\leq 3.9\%$, average distance ≤ 1.5 mm, root mean square (RMS) distance ≤ 2.8 mm and maximal distance ≤ 19.5 mm could be achieved. (2) The proposed method is highly efficient such that the overall segmentation can be finalized within 2 minutes.

Keywords: Segmentation, Kidney, Renal Cortex, Graph Cut.

1 Introduction

Kidney cancer is among the 10 most common cancers in both men and women. Overall, the lifetime risk for developing kidney cancer is about 1 in 70 (1.43%) [1]. Renal cell carcinoma which arises from the renal cortex is the most common type of kidney cancer in adults, responsible for approximately 80% of cases [2]. Hence, the investigation of the renal cortex has great value for kidney cancer research. Segmentation of the renal cortex is of particular clinical importance. Currently, many of the renal cortex segmentations used clinically rely mainly on manual methods, which are subjective, tedious, and prone to errors. Therefore there is a strong need to have a fully automatic and accurate kidney and renal cortex segmentation method in clinics.

Segmentation of renal cortex is not a trivial task. There were several prior investigations [3-15] in renal and renal cortex segmentation on CT and MRI images, including both semi-automatic [3-6] and fully automatic [7-9, 10-15] methods. For kidney segmentation in dynamic MR images, not only the spatial information but also the timing activities (also known as time intensity curves, TIC) were used for kidney segmentation [5, 10-12, 15]. Boykov et al. [11, 12] introduced a temporal Markov model to describe the TIC for each pixel, and used min-cut [11] for kidney segmentation. However, the user needs to specify seed points. For the atlas based and registration based methods, Sun et al. [13] presented an integrated image registration algorithm to segment renal cortex for dynamic renal perfusion MR images. Another registration based approach was proposed by Zollner et al. [14] by separating the inner compartments by k-means after a non-rigid registration guided by TIC. Song et al. [15] proposed a 4D level set framework for dynamic MR images kidney segmentation. The method also combined information from spatial anatomical structures and temporal dynamics.

Most of the studies described above [4, 5, 9-15] classified the kidney into the three compartments: renal cortex, medulla and pelvis. Based on the strict definition of renal cortex, only out-layer of the kidney is considered as a cortex because renal columns have anatomical and functional differences [16]. Many prior investigations [4, 9-11], [14] considered cortex and column as one tissue although they are anatomically different. However, in most clinical application for renal cortex segmentation, only the out-layer of the kidney is measured [16]. For certain clinical investigations, it is desirable to measure the cortex precisely. In this paper, we propose a framework to automatically segment renal cortex precisely.

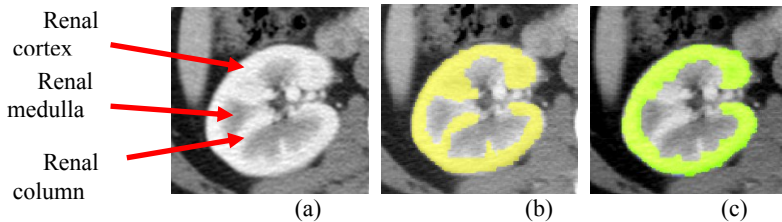


Fig. 1. Examples of segmentation results for cortex segmentation. (a) Input image; (b) Segmentation results by traditional GC method; (c) Segmentation results by the proposed method.

The proposed framework consists of two phases: training and segmentation phase. In the training phase, we construct the AAM model and train the live wire (LW) cost function and GC parameters. This allows us to construct oriented AAM (OAAM). The segmentation phase consists of two main parts: initialization and delineation. For the initialization part, a pseudo-3D initialization strategy is employed and a multi-resolution OAAM is proposed, which synergistically combines the AAM and LW methods (the essence of OAAM). A multi-resolution multi-object strategy is utilized to help for initialization. For the delineation part, the shape information generated from the initialization step is integrated into the GC cost computation, and an iterative algorithm is proposed for cortex delineation.

2 IGCOAAM Based Renal Cortex Segmentation

Fig. 2 shows the flowchart of the proposed method. In the training phase, an AAM is constructed and the LW and GC parameters are trained. The segmentation phase consists of two main steps: initialization and delineation. In the initialization step, we employ a pseudo 3D initialization strategy in which the kidney is initialized slice-by-slice via a multi-resolution multi-object OAAM method. The employment of pseudo-3D initialization strategy is motivated by two reasons: (1) Compared to a full 3D initialization method, the proposed method is much faster while achieved a similar performance, (2) It is difficult to combine AAM with LW in 3D. For the delineation part, the shape information generated from the OAAM initialization step is integrated into the GC cost computation. The kidney is segmented using the iterative graph cut oriented active appearance model (IGC-OAAM) method [17]. After getting the kidney contour, morphological operations are performed to get the initial cortex shape. Finally, the shape constrained IGC-OAAM method is applied once more to obtain the final renal cortex segmentation. The details of each step are given in the following sub-sections.

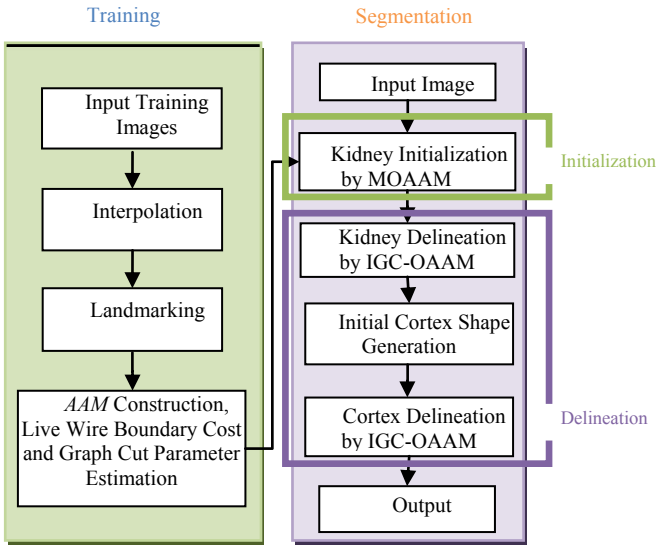


Fig. 2. The flowchart of the proposed system

2.1 Model Building and Parameter Training

The model building requires several constraints, one of which is anatomical correspondence finding. Since, the location and the size of the kidney may change considerably from subject to subject, it would be beneficial to interpolate the image slices of the organ in terms of having remarkable physical location correspondences between subjects. Therefore, prior to building the model, the top and bottom slices of kidney are first manually identified. Then linear interpolation is applied to generate

the same number of slices for each subject in training set. 2D OAAM models are then constructed for each slice level from the training images. Once the landmarks are specified, the standard AAM method [21] is used for constructing the model. Although we employ a pseudo-3-D initialization strategy, we also build the fully 3-D AAM, which is denoted as M_{3D} , using the method in [22].

The LW cost function and GC parameters are also estimated in this stage. For the training of LW cost function parameters, an oriented boundary cost function is devised as per the LW method [23]. For the GC parameters $P(I_p | O)$ and $P(I_p | B)$ (see equation (3)), the histograms of intensity for each object can also be estimated from the training images.

2.2 Initialization

The initialization step plays an important role in the overall framework. It does not only provide shape constraints to the later GC segmentation, but also makes the proposed method fully automatic. The proposed initialization method includes three main steps. First, a slice localization method is applied to detect the top and bottom slices of the kidney. Then a linear interpolation is applied to generate the same number of slices for the subject as in the model. Second, the organ is recognized slice by slice via a multi-resolution OAAM method. A multi-object strategy [17] is utilized to further assist the kidney initialization. The initialization performance with multiple organs in the model is much better than with a single organ due to the constraints among multiple organs [18]. The multi-resolution strategy is also used to improve the performance of the proposed method. Finally, a refinement method is applied to the initialization result.

2.3 Delineation by Shape Constrained GC

The proposed delineation method consists of three main steps. First, the kidney is segmented using the iterative shape constrained GC method which incorporates the kidney shape information generated from the initialization step. Second, after getting the kidney contour, morphological operations are performed to get the initial cortex shape. Finally, the shape constrained GC method is applied again to refine the final renal cortex which incorporates the cortex shape at this time. The details of each step are given in the following sub-sections. Since the shape constrained GC methods is the building block of the proposed delineation algorithm, we present it first.

2.3.1 Shape Constrained Graph Cut Cost Function

The segmentation problem can be formulated as an energy minimization problem such that for a set of pixels P and a set of labels L , the goal is to find a labeling $f: P \rightarrow L$ that minimizes the energy function $En(f)$.

$$En(f) = \sum_{p \in P} R_p(f_p) + \sum_{p \in P, q \in N_p} B_{p,q}(f_p, f_q), \quad (1)$$

where N_p is the set of pixels in the neighborhood of p , $R_p(f_p)$ is the cost of assigning label $f_p \in L$ to p , and $B_{p,q}(f_p, f_q)$ is the cost of assigning labels $f_p, f_q \in L$ to p and q . In our framework, the unary cost $R_p(f_p)$ was the sum of a data penalty $D_p(f_p)$ and a shape penalty $S_p(f_p)$ term. The data term was defined based on the image intensity and can be considered as a log likelihood of the image intensity for the target object. The shape prior term is independent of image information, and the boundary term is based on the gradient of the image intensity.

The proposed shape-integrated energy function was defined as follows:

$$En(f) = \sum_{p \in P} (\alpha \cdot D_p(f_p) + \beta \cdot S_p(f_p, \mathbf{x}_o)) + \sum_{p \in P, q \in N_p} \gamma \cdot B_{p,q}(f_p, f_q), \quad (2)$$

where α, β, γ are the weights for the data term, shape term, and boundary term, respectively, satisfying $\alpha + \beta + \gamma = 1$. These components are defined as follows:

$$D_p(f_p) = \begin{cases} -\ln P(I_p | O), & \text{if } f_p = \text{object label} \\ -\ln P(I_p | B), & \text{if } f_p = \text{background label} \end{cases} \quad (3)$$

$$B_{p,q}(f_p, f_q) = \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{d(p, q)} \cdot \delta(f_p, f_q), \quad (4)$$

$$\text{and } \delta(f_p, f_q) = \begin{cases} 1, & \text{if } f_p \neq f_q \\ 0, & \text{otherwise} \end{cases}, \quad (5)$$

where I_p is the intensity of pixel p , *object label* is the label of the object (foreground). $P(I_p | O)$ and $P(I_p | B)$ are the probability of intensity of pixel p belonging to object and background, respectively, which are estimated from object and background intensity histograms during the training phase (details given below). $d(p, q)$ is the Euclidian distance between pixels p and q , and σ is the standard deviation of the intensity differences of neighboring voxels along the boundary.

$$S_p(f_p, \mathbf{x}_o) = 1 - \exp\left(-\frac{d_{f_p}(p, \mathbf{x}_o)}{r_o}\right) \quad (6)$$

$$\text{and } d_{f_p}(p, \mathbf{x}_o) = \begin{cases} d(p, \mathbf{x}_o) & \text{if } f_p = \text{source label} \\ d(p, \bar{\mathbf{x}}_o) & \text{otherwise} \end{cases}, \quad (7)$$

where $d(p, \mathbf{x}_o)$ is the distance from pixel p to the set of pixels which constitute the interior of the current shape \mathbf{x}_o of object O . (Note that if p is in the interior of \mathbf{x}_o , then $d(p, \mathbf{x}_o) = 0$.) $d(p, \bar{\mathbf{x}}_o)$ is the distance from voxel p to the complementary of the shape \mathbf{x}_o . r_o is the radius of the sphere that roughly encloses \mathbf{x}_o . The linear time distance transform method in reference [18] was used in this paper for computing this distance.

2.3.2 Renal Cortex Segmentation

Now, we present the details for the renal cortex segmentation, which consists of three main steps: kidney delineation, cortex shape generation and cortex delineation.

Kidney Delineation. The purpose of this step is to precisely delineate the kidney shapes recognized in the previous step. We propose an IGC-OAAM method for the kidney's delineation, which effectively integrates the shape information from the OAAM initialization step with the globally optimal 3D delineation capability of the GC method. (see below algorithm).

We assume that the recognized shapes are sufficiently close to the actual boundaries in the given image to be segmented. The IGC-OAAM algorithm then determines what the new position of the landmarks of the objects represented in the initialized shape x_{in} should be such that the minimum GC cost is achieved, as presented below.

Algorithm: IGC-OAAM for Kidney Delineation

Input: Initialized shapes x_{in} , shape distance $d = \infty$, $\varepsilon =$ a small value.

Output: Resulting shapes x_{out} and the associated kidney boundaries shape KS.

Begin

While $d > \varepsilon$ do

1. Perform GC segmentation using equation (6) based on the OAAM initialized shapes x_{in} ;
2. Compute the new position of the landmarks by moving each landmark in x_{in} to the point closest on the GC boundary; call the resulting shapes x_{new} ;
3. If no landmarks move, then, set x_{new} as x_{out} and stop; Else, subject x_{new} to the constraints of model M_{3D} , and call the result x_{in} . And computer the shape distance d between the current shape and previous shape using the method in [19].

EndWhile

Perform one final GC segmentation based on x_{out} , and obtain the associated shape boundaries KS.

End

In our implementation, $\varepsilon = 0.1$. We found from the experimental results that the segmentation accuracies usually don't change much after two iterations. We also limit that a landmark can move within any iteration to 6 voxels in order to make the surface change smoother.

Renal Cortex Shape Construction. After getting the kidney contour, our target is to get the initial cortex shape. A heuristic method based on morphological operation (erosion) is used to construct an initial renal cortex shape. The number of erosion time is estimated from the training data set which is actually related with the depth of the renal cortex.

Renal Cortex Delineation. Based on the above step, we got the initial cortex shape. We assume that the initialized cortex shapes are sufficiently close to the actual boundaries in the given image to be segmented. Here, the IGC-OAAM is used again to refine the final renal cortex boundaries, and the generated initial cortex shape is used as the shape constraint.

3 Experimental Results

This study involved 27 patients (12 men and 15 women, age ranged from 19-63). Abdominal images were acquired during preoperative screening. All examinations were performed with one of two different types of CT scanners (LightSpeed Ultra, GE Medical Systems, Milwaukee, WI; or Mx8000 IDT 16, Philips Medical Systems, Andover, MA). Prior to image acquisition, the patients were injected with 130 mL of Isovue-300 contrast agent (Bracco Diagnostics, Milan, Italy). CT images were reviewed with a three-dimensional multiplanar reformatting interactive mode on an image-processing workstation (Advanced Workstation; GE Medical Systems). The pixel size varied from 0.55 to 1 mm, and slice thickness from 1 to 5 mm. The leave-one-out strategy was used to evaluate the proposed method.

For quantitative evaluation, the MICCAI 2007 grand challenge for liver segmentation evaluation criteria [19]: *volumetric overlap error*, *relative volume difference*, *average symmetric surface distance*, *root mean square (RMS) symmetric surface distance*, and *maximal symmetric surface distance* were used to evaluate the proposed method. The evaluation results are summarized in Table 1. The average volumetric overlap error for kidney segmentation is about 3.6%, while for cortex segmentation, the average overlap error is about 12.7%. In Fig. 3, the 3rd and 4th column shows the segmentation results for the kidney and cortex by the proposed method, respectively. Additionally, the error of cortex segmentation is higher than kidney segmentation, which may be due to the greater difficulty in cortex segmentation than kidney segmentation.

Table 1. Segmentation evaluation based on the “MICCAI 2007 grand challenge for liver segmentation” evaluation criteria

Method	Overlap Error [%]	Volume difference [%]	Avg. distance [mm]	RMS distance [mm]	Max distance [mm]
Kidney Segmentation	3.6 ± 2.6	1.9 ±2.1	0.9 ±0.5	1.7 ±1.1	16.9 ±7.1
Cortex Segmentation	12.7 ± 3.3	3.9 ±5.2	1.5 ±1.1	2.8 ±2.6	19.5 ±9.8

In terms of efficiency, Table 4 shows the computation time for the kidney and cortex segmentation on an Intel Xeon E5440 workstation with 2.83GHz CPU, 8 GB of RAM. For kidney segmentation, the time for the computation was reduced from 8 minutes for manual segmentation to about 1.3 minutes for automatic segmentation. For renal cortex segmentation, the time was reduced from 20 minutes for manual segmentation, to less than 2 minutes for automatic segmentation.

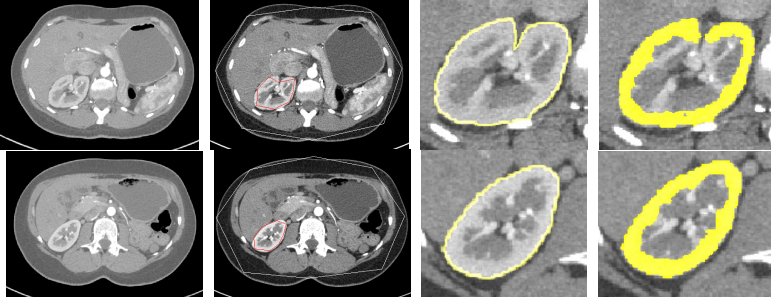


Fig. 3. Experimental results for two slice levels of kidney and cortex segmentation. The first column is the original slice image; the second is *MOAAM* initialization result; the third is the kidney segmentation results by the proposed method; the fourth is the cortex segmentation results by the proposed method.

4 Conclusions and Discussions

In this paper, we proposed a fully automated framework for renal cortex segmentation and quantification. The proposed framework synergistically combined the AAM, LW, and GC methods to exploit their complementary strengths. The proposed method consists of three main parts: model building, initialization, and segmentation. For the initialization part, we employed a pseudo-3D strategy, and segment the organs slice by slice via multi-resolution multi-object OAAM method, which effectively combines the AAM and LW methods. For the segmentation (delineation) part, an IGC-OAAM method was proposed which effectively integrated the shape information gathered from OAAM initialization with a GC algorithm. The method was tested on a clinical abdominal CT dataset with 27 contrast-enhanced images. The experimental results suggested that an overall renal cortex segmentation accuracy with overlap error $\leq 12.7\%$, volume difference $\leq 3.9\%$, average distance ≤ 1.5 mm, RMS distance ≤ 2.8 mm and maximal distance ≤ 19.5 mm could be achieved.

Use of the automatic algorithm greatly reduced the time needed for the segmentation, especially for cortex segmentation. The reduction of the mean segmentation time from 20 minutes for manual segmentation to less than 2 minutes for automatic segmentation made cortex segmentation from CTA images much more practically in clinical routine.

References

1. American Cancer Society,
<http://www.cancer.org/Cancer/KidneyCancer/DetailedGuide/kidney-cancer-adult-key-statistics>
2. Rini, B.I., Rathmell, W.K., Godley, P.: Renal Cell Carcinoma. *Curr. Opin. Oncol.* 20, 300–306 (2008)

3. de Priester, J.A., Kessels, A.G., Giele, E.L., den Boer, J.A., Christiaans, M.H., Hasman, A., van Engelshoven, J.M.: MR Renography by Semiautomated Image Analysis: Performance in Renal Transplant Recipients. *J. Magn. Reson. Imaging* 14, 134–140 (2001)
4. Shen, W., Kassim, A.A., Koh, H.K., Shuter, B.: Segmentation of Kidney Cortex in MRI Studies: A Constrained Morphological 3D H-Maxima Transform Approach. *Int. J. Med. Eng. Inform.* 1, 330–341 (2009)
5. Chevaillier, B., Ponvianne, Y., Collette, J.L., Claudon, M., Pietquin, O.: Functional Semi-Automated Segmentation of Renal DCE-MRI Sequences. In: 33th IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 525–528. IEEE Press, New York (2008)
6. Shim, H., Chang, S., Tao, C., Wang, J.H., Kaya, D., Bae, K.T.: Semiautomated Segmentation of Kidney from High-Resolution Multidetector Computed Tomography Images Using a Graph-Cuts Technique. *J. Comput. Assist. Tomogr.* 33, 893–901 (2009)
7. Freiman, M., Kronman, A., Esses, S.J., Joskowicz, L., Sosna, J.: Non-parametric Iterative Model Constraint Graph min-cut for Automatic Kidney Segmentation. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part III. LNCS, vol. 6363, pp. 73–80. Springer, Heidelberg (2010)
8. Lin, D.T., Lei, C.C., Hung, S.W.: Computer-Aided Kidney Segmentation on Abdominal CT Images. *IEEE Trans. Inf. Technol. Biomed.* 10, 59–65 (2006)
9. Ali, A.M., Farag, A.A., El-Baz, A.S.: Graph Cuts Framework for Kidney Segmentation with Prior Shape Constraints. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 384–392. Springer, Heidelberg (2007)
10. Tang, Y., Jackson, H.A., De Filippo, R.E., Nelson, M.D., Moats, R.A.: Automatic Renal Segmentation Applied in Pediatric MR Urography. *International Journal of Intelligent Information Processing (IJIIIP)* 1, 12–19 (2010)
11. Boykov, Y., Lee, V.S., Rusinek, H., Bansal, R.: 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)
12. Rusinek, H., Boykov, Y., Kaur, M., Wong, S., Bokacheva, L., Sajous, J.B., Huang, A.J., Heller, S., Lee, V.S.: Performance of an Automated Segmentation Algorithm for 3D MR Renography. *Magn. Reson. Med.* 57, 1159–1167 (2007)
13. Sun, Y., Jolly, M.P., Moura, J.M.F.: Integrated Registration of Dynamic Renal Perfusion MR Images. In: 11th IEEE International Conference on Image Processing, pp. 1923–1926. IEEE Press, New York (2004)
14. Zöllner, F.G., Sance, R., Rogelj, P., Ledesma-Carbayo, M.J., Rørvik, J., Santos, A., Lundervold, A.: Assessment of 3D DCE-MRI of the Kidneys Using Non-Rigid Image Registration and Segmentation of Voxel Time Courses. *Comput. Med. Imaging Graph.* 33, 171–181 (2009)
15. Song, T., Lee, V.S., Rusinek, H., Bokacheva, L., Laine, A.: Segmentation of 4D MR Renography Images Using Temporal Dynamics in a Level Set Framework. In: 5th IEEE International Symposium on Biomedical Imaging, pp. 37–40. IEEE Press, New York (2008)
16. Padigala, K., Hartle, J.E., Kirchner, H.L., Schultz, M.F.: Renal Cortical Thickness as a Predictor of Renal Function and Blood Pressure Status Post Renal Artery Stenting. *Angiology* 60, 719–724 (2010)
17. Chen, X., Udupa, J.K., Alavi, A., Torigian, D.A.: Automatic Anatomy Recognition via Multiobject Oriented Active Shape Models. *Med. Phys.* 37, 6390–6401 (2010)
18. Ciesielski, K.C., Chen, X., Udupa, J.K., Grevera, G.J.: Linear Time Algorithms for Exact Distance Transform. *J. Math. Imaging Vis.* 39, 193–209 (2011)

19. Heimann, T., van Ginneken, B., Styner, M.A., et al.: Comparison and Evaluation of Methods for Liver Segmentation from CT Datasets. *IEEE Trans. Med. Imaging* 28, 1251–1265 (2009)
20. Cootes, T.F., Edwards, G., Taylor, C.: Active Appearance Models. *IEEE Trans. Pattern Anal. Mach. Intell.* 23, 681–685 (2001)
21. Mitchell, S.C., Bosch, J.G., Lelieveldt, B.P., van der Geest, R.J., Reiber, J.H., Sonka, M.: 3-D Active Appearance Models: Segmentation of Cardiac MR and Ultrasound Images. *IEEE Trans. Med. Imaging* 21, 1167–1178 (2002)
22. Falcao, A.X., Udupa, J.K., Samarasekera, S., Sharma, S.: User-Steered Image Segmentation Paradigms: Live Wire and Live Lane. *Graph. Models Image Process.* 60, 233–260 (1998)

An Implicit Inter-subject Shape Driven Image Deformation Model for Prostate Motion Estimation

Bartłomiej W. Papież¹, Bogdan J. Matuszewski¹, Lik-Kwan Shark¹,
and Christopher Moore^{2,3}

¹ Applied Digital Signal and Image Processing Research Centre,
University of Central Lancashire, Preston, UK
`bartlomiej.papiez@eng.ox.ac.uk`

² Christie Medical Physics and Engineering,
The Christie NHS Foundation Trust, Manchester, UK

³ Faculty of Medical and Human Sciences,
Manchester Academic Health Sciences Centre,
University of Manchester, Manchester, UK

Abstract. This paper describes a novel approach for model based estimation of a dense deformation field utilizing an implicit representation of shape changes. Unlike existing methods based on the Point Distribution Model (PDM), the proposed method is not affected by an incorrect point correspondence which is a major limiting factor in practical applications of the PDM with clinical data. The proposed method uses regression between parametric representations of pelvic organs' shape and corresponding dense displacement field parameterized by the stationary vector field. The regression function is learned based on the training data sets including subjects with representative organ deformations, where the inter- and intra- subject correspondences are established via the log-Euclidean diffeomorphic formulation. The evaluation of the proposed method is conducted both on synthetic examples to provide systematic experimental evidence of correctness of the implicit shape representation for shape-driven prediction of the deformation field and, real MRI data to show accuracy in terms of deformation and prostate position prediction. The results show an increased robustness of the proposed framework in comparison to PDM approaches and suggest potential of its application for adaptive radiation therapy of prostate.

Keywords: Prostate motion estimation, deformation modeling, diffeomorphic image registration, implicit shape representation, regression.

1 Introduction

The estimation of a reliable deformation field from noisy data has essential importance in many clinical applications. In some cases when it is possible to obtain high quality training data, a deformation (motion) model constructed from such

data, can be either introduced into the registration process as an additional constraint [1] or linked somehow to partial information from a measurement that is affected by the deforming tissue and thus used to indirectly infer the tissue deformation [20,9,8,4].

Currently, such measurement is most often restricted to a point or a set of points. Implementations of this methodology include online systems for simultaneous compensation of laparoscopic camera motion and respiration induced tissue deformations [11], 4D image-guided lung radiation therapy [20,9,8], pelvic area subject-specific [17] and population based [16] organs modeling. The shape variability of rectum and bladder with applications to radiotherapy planning and delivery was also considered in [14], where the statistical deformation model was built on parameters of spherical harmonics used to represent organs surface with a small number of points interactively selected by an operator. For the purpose of lung motion estimation, the air content of a reference volume as a predictor has been used [4]. All the aforementioned methods require the estimation of correspondence between points detected in images representing different deformation phases, which is sensitive to occurrences of gross error (data outliers). To avoid the problems associated with point correspondence errors, it is proposed here to use an implicit representation for shape description. The implicit shape representation was considered and successfully utilized for segmentation of medical images containing known types of objects [18], and also in a global alignment and local registration of shapes [6]. In both cases, this representation solves difficulties with finding correspondences.

Whereas the intra-patient motion model can be easily generated utilizing multiple registrations between acquired volumes representing different phases of organs' motion [17,20], the inter-patient motion model that is built in similar manner, is biased towards the chosen reference image. Therefore, a *mean* reference patient has to be estimated followed by establishment of the *common patient space* [5]. Moreover, in many recent approaches ([20,9,8]), the statistical models of the deformation fields were obtained by directly applying principal components analysis (PCA) to them (or B-spline parameterization [15]). Although those approaches were to some extent successful in certain applications, this motion model has some disadvantages. The most important is that such eigenvectors do not form a vector space of valid deformation fields (e.g. diffeomorphism is not guaranteed to be preserved). Here, the log-Euclidean framework [2,3,19] is applied to parameterize the non-linear diffeomorphic deformation field.

The main contribution of this paper is a novel approach to prostate position and dense deformation field prediction. The proposed method uses implicit surfaces (described in Section 2.3) to represent the shape of bladder and rectum. The results presented in Section 3 show the advantages of this shape descriptor in comparison to PDM. Moreover, the *common patient space* that is essential for the presented method and established in log-Euclidean framework by performing group-wise registration, is summarized in Section 2.1. The intra-subject registration to estimate subject-specific deformation is explained in Section 2.2. Finally, the training process is demonstrated in Section 2.4. The results obtained

for the simulated and real magnetic resonance imaging (MRI) data are shown in Section 3 to demonstrate the plausibility of the described prostate position prediction system. The results for synthetic data illustrate efficiency of the proposed method in comparison to PDM. The accuracy with respect to prostate position estimation based on real MRI data are shown to demonstrate the potential application in ART.

2 Methodology

In the proposed approach, each patient $k : k = 1, \dots, K$ is described by a set of n -dimensional images $\mathbf{I}^k = \{I_l^k : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}, l = 0, \dots, L^k\}$. A transformation $\varphi_{0l} : \Omega \rightarrow \Omega$ that warps image I_l^k to image I_0^k , is represented by a stationary velocity field \mathbf{v}_{0l}^k , related to a diffeomorphic displacement field \mathbf{u}_{0l}^k through the exponential mapping $\varphi_{0l}^k(\mathbf{x}) = \mathbf{x} + \mathbf{u}_{0l}^k(\mathbf{x}) = \mathbf{x} + \exp(\mathbf{v}_{0l}^k(\mathbf{x}))$ [19]. The diffeomorphic registration guarantees that folding will not take place, contrary to many other methods which do not have this constraint included [16,20].

2.1 Common Space Generation

Calculating the velocity fields for each set of patient images can only provide a patient specific motion model. In order to obtain a general motion model, the *common patient space* has to be established. To avoid possible bias due to selecting a particular reference image, the implicit group-wise image registration was proposed in [5]. Contrary to the original method, in this work to maintain diffeomorphic consistency, the displacement fields are parameterized via stationary velocity fields. The energy function $\varepsilon(\mathbf{v})$ is given in Equation 1 where *Sim* is a similarity measure and *Reg* is a regularization term, and the Demon-like approach is used to minimize it [12].

$$\varepsilon(\mathbf{v}) = \sum_{k_1=1}^K \sum_{\substack{k_2=1 \\ k_1 \neq k_2}}^K \text{Sim}(I_0^{k_1}(\exp(\mathbf{v}_0^{k_1})), I_0^{k_2}(\exp(\mathbf{v}_0^{k_2}))) + \alpha \sum_{k_1=1}^K \text{Reg}(\mathbf{v}_0^{k_1}) \quad (1)$$

The implicit group-wise registration estimates the set of transformations φ_{0m}^k that warps each planning image I_0^k to the *mean* intensity image I_m , that finally can be calculated as an average of all warped images $\frac{1}{K} \sum I_0^k(\varphi_{0m}^k)$.

The performance of the proposed method is demonstrated in Fig. 1, showing the average of the input data and the Intensity Variance (IV) [5,12] before and after data co-registration.

2.2 Intra-subject Registration

Generation of the deformation model requires estimation of numerous non-linear diffeomorphic transformations between different stages of organs motion. Since the transformations φ_{0m}^k have already been established, the intra-patient deformations are calculated for the images $J_l^k = I_l^k(\varphi_{0m}^k)$ in the *common patient*

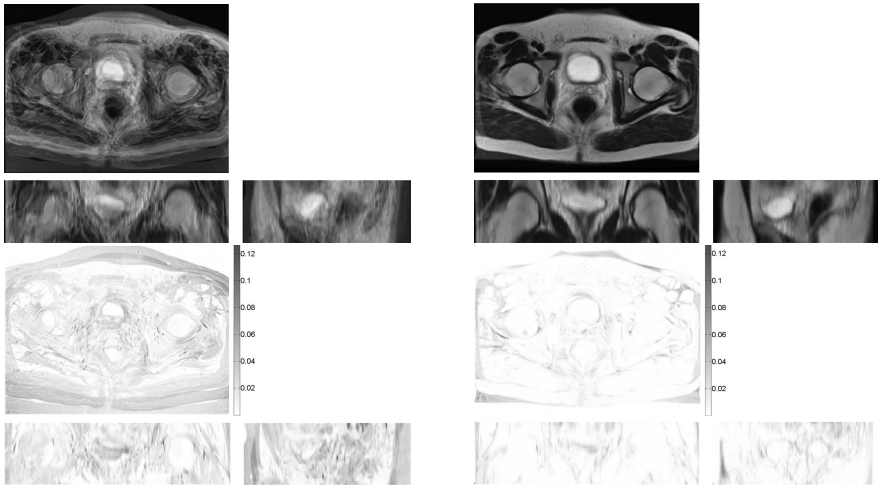


Fig. 1. Results of performing the log-domain implicit group-wise registration for the MRI pelvic-area data set: (top left) average of the input data; (bottom left) IV before registration; (top right) average of the input data after performing registration; (bottom right) IV after registration

space. To assure symmetry of the registration with respect to the input images and impose the inverse consistency criterion, the consistent intensity-based algorithm previously proposed in [19] is used. This approach implements a fast and efficient Demon method, that minimizes the energy:

$$\varepsilon(\mathbf{v}_{0l}^k) = Sim(J_0^k, J_l^k(\exp(\mathbf{v}_{0l}))) + Sim(J_l^k, J_0^k(\exp(-\mathbf{v}_{0l}))) + \alpha Reg(\mathbf{v}_{0l}) \quad (2)$$

The main advantage of using this method of registration is that the output velocity field \mathbf{v}_{0l} minimizing energy ε (Equation 2) can be directly used for a log-domain calculation of deformation statistics without explicitly computing the logarithm of the displacement field, which is time consuming [23]. However, different methods of registration can be used instead as long as they permit *vectorial* statistics to be calculated and the invertibility constraint is preserved [2].

2.3 Shape Representation

To evaluate the proposed surrogate-based image deformation prediction methodology, segmentation of an n -dimensional image J_l^k has to be performed to produce a binary image $S_l^k : \Omega \subset \mathbb{R}^n \rightarrow \{0, 1\}$ representing the organ of interest. Segmentation for organs such as bladder or rectum is a relatively simple task when compared to prostate segmentation. Such segmentation can often be achieved using one of the previously proposed automatic or semi-automatic segmentation algorithms [21]. The sign distance function $\phi_l^k(\mathbf{x})$ that is chosen here as a shape descriptor, is defined as the minimum Euclidean distance from a given

voxel position $\mathbf{x} = [x^1, \dots, x^n] \in \Omega$ to the shape's boundary S_l^k , multiplied by 1 or -1 depending on which side of the boundary \mathbf{x} is located. The positive or negative distance is assigned to all voxels inside or outside S_l^k respectively, whereas for all voxels on the shape boundary the distance function has value zero. An efficient algorithm, with linear computation time, for exact Euclidean Distance (ED) calculation for arbitrary dimensional binary image was proposed in [10].

2.4 Training Process

During the training process, the motion model, linking observed shape changes of the bladder/rectum and corresponding deformation field variations of the pelvic region, is estimated using data sets of implicit shapes and matched velocity fields. First, note that due to the high dimensionality of the velocity field and adopted shape representation, a direct estimation of the motion model is computationally prohibitive. A commonly used technique for reducing dimensionality of data is PCA [20,9,8]. In our case, it allows the extraction of the major modes of variation for shape and velocity fields. For each registered image J_p ($p = 1, \dots, P$), where P is the number of volumes in the database, different from the reference volume $P = K \cdot (L - 1)$) the corresponding velocity field and implicit shapes are formed as vectors. After carrying out PCA using all V_p , the velocity field can be approximated as:

$$V_p \simeq \mu_V + \sum_{i=1}^{E_V} \gamma_{pi} \mathbf{w}_i = \mu_V + \Gamma_p W_V \quad (3)$$

where μ_V is the sample mean velocity field, \mathbf{w}_i is i th eigenvector, γ_{pi} are coefficients, W_V is the matrix of eigenvectors and Γ_p is the vector of coefficients with E_V representing the number of the major velocity field eigenmodes observed in the training data set. Similarly, for the implicit shape representation:

$$\Phi_p \simeq \mu_\Phi + \sum_{i=1}^{E_\Phi} \theta_{pi} \mathbf{z}_i = \mu_\Phi + \Theta_p Z_\Phi \quad (4)$$

where μ_Φ is the sample mean implicit shape, \mathbf{z}_i is i th eigenvector, θ_{pi} are coefficients, Z_Φ is the matrix of eigenvectors and Θ_i is the vector of coefficients with E_Φ representing the number of the major implicit shape eigenmodes present in the training data set. The sign distance functions do not form a vector space and although it is possible to "vectorise" this representation by using so called Log-Odds maps [13] in the proposed algorithm no operations are performed in that space, since the θ_{pi} coefficients are only used in regression model to predict velocity field. Additionally it was shown in [18] that this theoretical inconsistency does not affect significantly the performance of this shape representation in practice. The motion model M , linking observed organ deformations with the corresponding velocity fields is built via multivariate linear regression between coefficients matrixes $\Gamma = [\Gamma_0^T, \dots, \Gamma_p^T, \dots, \Gamma_P^T]^T$ and $\Theta = [\Theta_0^T, \dots, \Theta_p^T, \dots, \Theta_P^T]^T$:

$$M = (\Theta^T \Theta + \rho D)^{-1} \Theta^T \Gamma \quad (5)$$

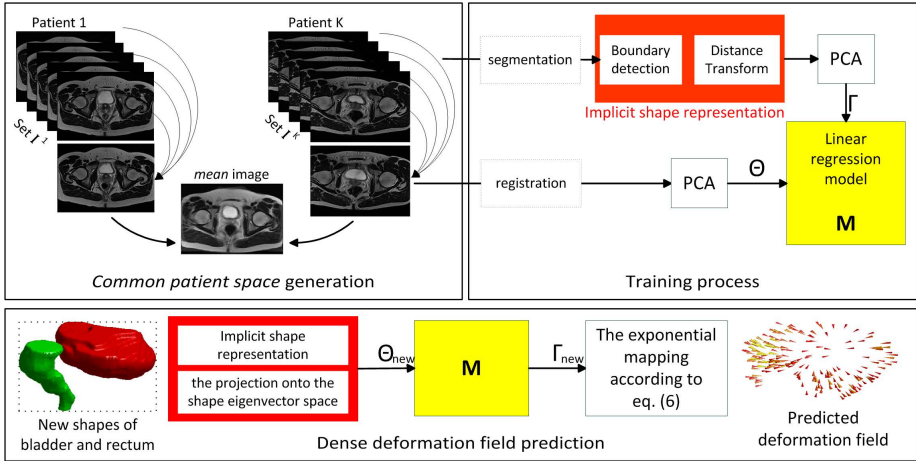


Fig. 2. Training process and deformation field prediction using the implicit shape-based dense deformation field estimation model for prostate

where ρ denotes a regularisation parameter for the prediction model and D is an identity matrix [16].

2.5 Dense Deformation Field Prediction

Once the model M is learned, it can be utilized in practice both on the images for patients already included in the model and for images of new patients. In both cases, a new image has to be first mapped to the *common patient space* and then segmented to obtain its shape descriptor. Then, the implicit representation of any new shape Φ_{new} can be directly projected onto the shape eigenvector space Z_Φ to get shape coefficients Θ_{new} . Subsequently, corresponding velocity coefficients Γ_{new} can be calculated using the estimated motion model: $\Gamma_{new} = \Theta_{new}M$. Then, under the Baker-Campbell-Hausdorff formula for the composition of the diffeomorphic deformations [3], the predicted diffeomorphic deformation field u_{new} can be approximated by exponentiation of the sum of the velocity fields eigenvectors:

$$u_{new} = exp(\mu_V + \Gamma_{new}W_V) \tag{6}$$

The computational burden of this training scheme is highly time-consuming, mostly due to multiple registration of volumes, but it is performed only once during the learning of the motion model M . However, the model evaluation is fast as it involves only the calculation of implicit shape descriptors, matrix multiplication and calculation of the exponential map. The scheme presenting the proposed model training framework and the corresponding model evaluation process is shown in Fig. 2.

Table 1. Errors of estimation deformation fields obtained using different shape representations: (PDM perf.) - PDM with ground truth correspondence; (PDM norm.) - PDM with Gaussian noise $\mu = 0.0$, $\sigma = 1.0$ representing typical measurement error; (PDM g1.) - PDM norm. with gross error $\zeta = 5\%$, τ_{10}^{20} ; (PDM g2.) - PDM norm. with gross error $\zeta = 5\%$, τ_{20}^{40} ; (PDM g3.) - PDM norm. with gross error $\zeta = 1\%$, τ_{20}^{40} . All results are given in *mm*.

	Trial 1		Trial 2		Trial 3		Trial 4		Trial 5		avg \pm std
	avg	max	avg	max	avg	max	avg	max	avg	max	
PDM perf.	0.23	3.61	0.25	3.90	0.21	3.36	0.18	2.75	0.18	2.7	0.21 \pm 0.12
PDM norm.	0.47	4.94	0.52	6.63	0.49	4.82	0.44	4.46	0.44	3.83	0.47 \pm 0.11
PDM g1.	0.79	8.77	0.72	7.26	0.76	12.04	0.60	6.52	0.65	6.26	0.70 \pm 0.22
PDM g2.	1.19	17.59	1.24	13.11	1.14	12.18	1.07	16.45	1.06	10.16	1.14 \pm 0.34
PDM g3.	0.63	6.08	0.59	7.22	0.64	10.17	0.54	5.81	0.67	7.71	0.61 \pm 0.22
implicit	0.31	5.36	0.36	4.14	0.30	5.34	0.25	3.42	0.28	4.13	0.30 \pm 0.14

3 Experimental Results

The simulated data set consists of 50 volumes of size 320x240x30, which were warped by known (ground truth) deformation fields. The known velocity fields were generated from a random zero-mean Gaussian *velocity process* defined on selected knots of a sparse 3D regular grid of size 30x20x15, which were heavily smoothed by a low-pass filter. The selected knots were chosen in the area close to the bladder. Then, velocity fields were interpolated to a finer grid with cubic splines before being scaled and exponentiated to get a wide range of diffeomorphic deformation fields [3]. The resultant data set was divided into two parts. The first set with 40 volumes was used for initializing the motion model, whereas 10 remaining volumes were used for model evaluation. The experiments were repeated 5 times, each time for a different, randomly selected training data set. For all the tests, a segmented bladder was used as a surrogate, represented either by the signed distance function for the method proposed in this paper or by 167 points uniformly distributed on the bladder surface for the PDM approach. To model the occurrence of incorrect point-to-point correspondences, gross errors were introduced to the PDM representation during the methodological validation. The implemented gross errors are characterized by: ζ - representing expected percentage of points affected by the gross errors and, τ_{min}^{max} - encoding the minimum (*min*) and maximum (*max*) magnitude of the gross errors, with the actual simulated error magnitudes selected randomly from a uniform distribution. In all cases the points were selected on the bladder surface.

The results shown in Tab. II indicate that the implicit shape representation performs in a similar manner to PDM without gross errors. PDM accuracy deteriorates quite quickly with even small levels of gross errors (with only one or two points affected out of 167 used for shape representation) thereby showing that the estimated deformation field is no longer reliable. As might be anticipated the estimation error gets bigger with the increasing magnitude of gross errors τ or when more points are affected by it (for larger ζ).

Table 2. Errors of estimation deformation fields and prostate positions

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 1-5 avg \pm std (<i>mm</i>)
Deformation characteristics	7.40 \pm 3.39	4.07 \pm 0.84	8.61 \pm 3.03	5.94 \pm 2.42	3.57 \pm 1.39	5.92 \pm 2.14
Error of prediction	2.45 \pm 1.16	1.45 \pm 0.15	2.71 \pm 0.98	2.05 \pm 0.78	1.25 \pm 0.41	1.98 \pm 0.62

For the second part of the experiments, real MRI data were used. The data set consisted of 25 scans obtained from 5 subjects. For each subject the images were collected in equal intervals over 40 minutes in order to capture significant changes of bladder/rectum shape. Then the images were normalized to the spatial resolution of $1.0 \times 1.0 \times 3.0 \text{ mm}$ and the anatomical structures used in experiments were manually annotated. For each experiment, four sets were chosen as training data, the remaining one as test data. The experiments were repeated 5 times, every time with a different set selected as test data. In order to get quantitative results and to estimate ground truth deformation fields, the registration was independently performed between corresponding pairs of images that were then used to compare with those estimated using the prediction system. The overall robustness in terms of the accuracy with respect to the ground truth deformation fields is summarized in the 2nd row of Tab. 2 whereas the 1st row characterizes the ground truth data. Based on these results, it can be concluded that the proposed method is able to estimate dense deformation field with errors less than $2.0 \pm 0.62 \text{ mm}$. Fig. 3 (left) shows the *mean* image with segmented bladder, prostate and rectum whereas Fig. 3 (middle) shows a new image with the actual position of the three organs alongside the superimposed position of the prostate from the *mean* image. The new position of the prostate from the *mean* image is estimated by warping the segmented *mean* image with the displacement field calculated based on the learned model and the measured shape of the bladder and rectum in the new image. The result shown in Fig. 3 (right) demonstrates that the proposed method, in this case, was able to estimate correctly the actual position of the prostate in the new data. Additionally, the computational cost is significantly reduced when compared with the classical registration. The average time of prediction when the motion model is already trained is about 1s while the registration takes about 180s.

4 Conclusion

The paper describes a novel technique for model-based image dense deformation field estimation with an implicit surface representation, with initial results suggesting this may be used as an effective and robust deformation descriptor. The proposed framework uses a motion model estimated from a training data set of shapes and corresponding displacement fields parameterized via stationary velocity fields estimated using a fast and efficient diffeomorphic registration

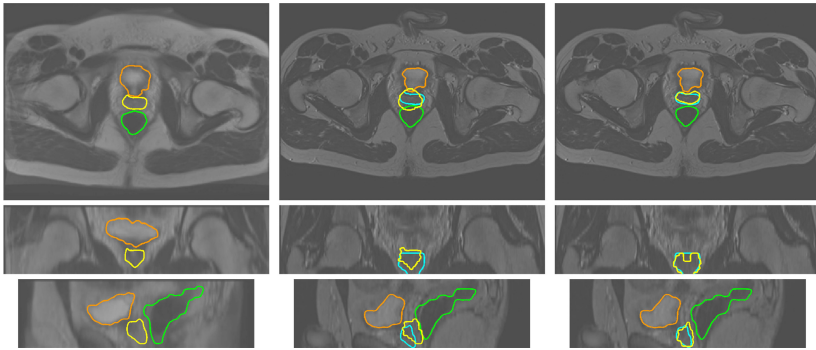


Fig. 3. (Left) *mean* image used during training with delineated bladder (magenta), rectum (green) and prostate (yellow). (Middle) Image, not used during model training, with delineated bladder, rectum and (ground truth) prostate in cyan. (Right) Image with the position of the prostate estimated using the proposed method (yellow).

scheme, formulated in the log-Euclidean framework. Additionally, it has been demonstrated that with the help of the proposed method, it may well be possible to predict the dense displacement field solely from the measured deformations of the implicit surface. Experiments conducted with the real data show that it is possible to predict the real-world deformation field, which suggests that the position of the prostate might be inferred from shape deformations of the bladder/rectum, which lie immediately superior and posterior respectively. Here we note that the relatively solid prostate surrounds the urethra as it emerges from the base of the bladder. Furthermore, prostatic motion is limited by lateral ligaments, which largely constrain its motion to lie in the anterior-posterior plane. Hence, bladder and rectal deformations due to filling are the primary drivers of prostate displacement. The seminal vesicles attached to the prostate gland are more challenging, due to their small size and deformable nature. Nevertheless, as it is relatively easier to segment bladder, somewhat less so the rectum, compared to the prostate from the X-ray cone beam CT scans now commonly deployed for image guided radiation therapy, it can be concluded that the proposed methodology can be potentially useful for adaptive radiation therapy of prostate. Thus, further investigation will be centered on combining the motion model built from pre-radiotherapy MRI data with shape descriptors extracted from radiotherapy imaging. Furthermore, it is worth noting, that a new generation of radiotherapy machines is being developed to provide in-room 3D live MRI concurrently with therapeutic radiation [7].

Acknowledgements. This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/D077540/1].

References

1. Albrecht, T., Lüthi, M., Vetter, T.: A statistical deformation prior for non-rigid image and shape registration. In: Proc. IEEE CVPR 2008, pp. 1–8 (2008)
2. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A Log-Euclidean Framework for Statistics on Diffeomorphisms. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006, Part I. LNCS, vol. 4190, pp. 924–931. Springer, Heidelberg (2006)
3. Bossa, M., Hernandez, M., Olmos, S.: Contributions to 3D Diffeomorphic Atlas Estimation: Application to Brain Images. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 667–674. Springer, Heidelberg (2007)
4. Ehrhardt, J., Werner, R., Schmidt-Richberg, A., Handels, H.: Statistical modeling of 4d respiratory lung motion using diffeomorphic image registration. *IEEE Trans. Med. Imag.* 30(2), 251–265 (2011), <http://dx.doi.org/10.1109/TMI.2010.2076299>
5. Geng, X., Christensen, G., Gu, H., Ross, T., Yang, Y.: Implicit reference-based group-wise image registration and its application to structural and functional mri. *Neuroimage* 47(4), 1341–1351 (2009), <http://dx.doi.org/10.1016/j.neuroimage.2009.04.024>
6. Huang, X., Paragios, N., Metaxas, D.: Shape registration in implicit spaces using information theory and free form deformations. *IEEE T-PAMI* 28(8), 1303 (2006)
7. Lambert, J., Greer, P., Menk, F., Patterson, J., et al.: MRI-guided prostate radiation therapy planning: Investigation of dosimetric accuracy of MRI-based dose planning. *Radiother. Oncol.* 98(3), 330–334 (2011), <http://dx.doi.org/10.1016/j.radonc.2011.01.012>
8. Liu, X., Davis, B., Niethammer, M., Pizer, S., Mageras, G.: Prediction-driven respiratory motion atlas formation for 4d image- guided radiation therapy in lung. In: MICCAI 2010 Pulmonary Image Analysis Workshop (2010)
9. Liu, X., Saboo, R., Pizer, S., Mageras, G.: A shape-navigated image deformation model for 4d lung respiratory motion estimation. In: Proc. ISBI, pp. 875–878 (2009), <http://dx.doi.org/10.1109/ISBI.2009.5193192>
10. Maurer, C., Qi, R., Raghavan, V.: A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *IEEE T-PAMI* 25(2), 265–270 (2003)
11. Mountney, P., Yang, G.-Z.: Motion Compensated SLAM for Image Guided Surgery. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part II. LNCS, vol. 6362, pp. 496–504. Springer, Heidelberg (2010)
12. Papiez, B.W., Matuszewski, B.J., Shark, L.K., Quan, W.: Facial expression recognition using log-euclidean statistical shape models. In: Proc. ICPRAM (2012)
13. Pohl, K.M., Fisher, J., Shenton, M.E., McCarley, R.W., Grimson, W.E.L., Kikinis, R., Wells, W.M.: Logarithm Odds Maps for Shape Representation. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006, Part II. LNCS, vol. 4191, pp. 955–963. Springer, Heidelberg (2006)
14. Price, G., Moore, C.: Comparative evaluation of a novel 3d segmentation algorithm on in-treatment radiotherapy cone beam ct images. In: Proceedings of the SPIE Conference on Medical Imaging, vol. 6512(3), pp. 38.1–38.11 (2007)
15. Rueckert, D., Frangi, A., Schnabel, J.: Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration. *IEEE Trans. Med. Imag.* 22(8), 1014–1025 (2003), <http://dx.doi.org/10.1109/TMI.2003.815865>

16. Shi, Y., Liao, S., Shen, D.: Learning statistical correlation for fast prostate registration in image-guided radiotherapy. *Med. Phys.* 38(11), 5980–5991 (2011), <http://dx.doi.org/10.1118/1.3641645>
17. Söhn, M., Birkner, M., Yan, D., Alber, M.: Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation. *Phys. Med. Biol.* 50(24), 5893–5908 (2005), <http://dx.doi.org/10.1088/0031-9155/50/24/009>
18. Tsai, A., Yezzi Jr., A., Wells, W., Tempny, C., Tucker, D., Fan, A., Grimson, W.E., Willisky, A.: A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. Med. Imag.* 22(2), 137–154 (2003)
19. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Symmetric Log-Domain Diffeomorphic Registration: A Demons-Based Approach. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I*. LNCS, vol. 5241, pp. 754–761. Springer, Heidelberg (2008), http://dx.doi.org/10.1007/978-3-540-85988-8_90
20. Zhang, Q., Pevsner, A., Hertanto, A., Hu, Y., Rosenzweig, K., Ling, C., Mageras, G.: A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Med. Phys.* 34(12), 4772–4781 (2007)
21. Zhang, Y., Matuszewski, B., Histace, A., Precioso, F., Kilgallon, J., Moore, C.: Boundary delineation in prostate imaging using active contour segmentation method with interactively defined object regions. In: *MICCAI 2010 Prostate Cancer Imaging Workshop*. pp. 131–142 (2010), <http://portal.acm.org/citation.cfm?id=1889234.1889249>

Segmentation of the Cartilage in the Rib Cage in 3D MRI

Yolanda H. Noorda, Lambertus W. Bartels, and Josien P.W. Pluim

Image Sciences Institute, UMC Utrecht, 3584 CX Utrecht, The Netherlands
Y.H.Noorda@umcutrecht.nl

Abstract. Interventional non-invasive MR-guided techniques for treatment of liver tumors, such as HIFU, could benefit greatly from automatic cartilage detection. In this paper, segmentation of the cartilage in the rib cage is performed in 3D MR images. This is a challenging task, due to the poor contrast between cartilage and muscle, and the non-uniform intensity of the cartilage.

Our segmentation algorithm is based on feature selection by analyzing orientation and vesselness, automatic sternum localization using anatomical knowledge, skeletonization and ridge finding, and level set evolution.

We show that our algorithm is capable of detecting all visible cartilage structures in the scans. Gaps and false positives may occur, due to lack of contrast or the presence of non-cartilage structures with similar features. However, the segmentation is accurate, even for regions with low contrast, with an average error of the boundary of 1.1 mm.

Keywords: Cartilage segmentation, rib cage, MRI.

1 Introduction

Cartilage segmentation of the rib cage in MRI is a relatively unexplored topic, since it has never served a purpose. Nowadays, new techniques for treatment of tumors in abdominal organs are in development, that require automatic monitoring of the treatment. MR-guided HIFU (High Intensity Focused Ultrasound) is an example of such a technique. An ultrasound transducer is used to create a heating focus at a tumor, such that the tumor tissue coagulates. The temperature is monitored by MR-thermometry. To use this technique for ablation of tumors in the liver, the ultrasound beam needs to propagate through the intercostal space, to prevent painful heating of the ribs. Therefore, the location of the ribs should be known during treatment. Since the liver is partially covered by the cartilage of the rib cage, automatic cartilage segmentation in MR images is required.

The rib cage consists of bone and cartilage. The first seven ribs are attached to the sternum by the costal cartilage. The 8th, 9th, and 10th ribs join with the costal cartilage of the 7th rib. The floating ribs are not attached to the costal cartilage.

MRI-scans can depict the costal cartilage to a certain extent, however, the intensity values are fairly inhomogeneous and in addition, low contrast complicates the detection.

There are no methods in literature yet which describe costal cartilage segmentation. However, several segmentation schemes exist for cartilage in the knee. Since the intensity of cartilage in MRI is non-uniform, intensity-based approaches will not solve the problem. Knee cartilage segmentation methods often include prior knowledge [1], classifiers [2,3] or a graph optimization framework [4].

Rib cage segmentation is often done on CT images, where model-based approaches or classifiers are used frequently [5,6,7].

In this paper, a fully automatic costal cartilage segmentation scheme is developed and evaluated.

2 Materials and Methods

2.1 Data

Since HIFU treatment of the liver is not yet applied in the clinic, there are no clinical data available which depict the costal cartilage on MRI. Therefore, MR images of the cartilage were obtained on four volunteers. However, there will be no anatomical differences with patients, since the pathology will be in the liver, and not in the cartilage.

In accordance with local regulations in the hospital, the volunteers were screened for contraindications and informed about possible risks.

For each volunteer, a 3D T1-weighted gradient echo coronal fat-suppressed MRI scan was acquired (Philips Achieva, TE/TR 1.940/4.017 ms, flip angle 10° , in-plane resolution $1.302 \times 1.302 \text{ mm}^2$ or $1.372 \times 1.372 \text{ mm}^2$, slice thickness 1.5 mm). The scan consisted of 67 slices, such that only the anterior part of the body was contained in the field of view. The volunteers were placed in prone position, to simulate HIFU-treatment. The images were acquired during normal breathing.

For validation, a sagittal scan was acquired for two volunteers using the same sequence as mentioned above. This scan consisted of 50 slices and was angulated in a way that the profile of the cartilage close to the liver is clearly visible, such that the accuracy of the segmentation boundaries can be evaluated on it.

2.2 Methods

For all images, a volume of interest containing the body is created by histogram analysis, thresholding and morphology.

The cartilage segmentation pipeline consists of five main steps: analyzing vesselness and orientation, sternum localization, preliminary segmentation of all structures connected to the sternum, skeletonization and ridge finding, and final segmentation by conditional growing using level sets. Below, each step will be explained in detail.

Vesselness and Orientation. One of the main features of the costal cartilage structures is their vessel-like shape. Therefore, as a first step, 3D vesselness is calculated on the image using the vesselness filter as proposed in [8]. The parameter σ , which indicates the scale, is varied from 3 to 7 mm in steps of 0.5 mm. After normalization, the vesselness image is thresholded at 5% to obtain a binary volume containing vessel-like structures. Fig. 1(a) and Fig. 1(b) show the original volume and the binary vesselness volume.

For each component of this binary vesselness volume, the orientation of every voxel is determined in-plane. Then every slice of each component is labeled with the orientation that occurs most often. This slice-by-slice approach makes it possible to disconnect structures that have a different in-plane orientation, such that cartilage structures will be separated from vessel-like structures in the liver and the intestines.

Finding a Point in the Sternum. The next step is to find a point (x_s, y_s, z_s) in the sternum from which the cartilage structures can be located. Since the sternum is not visible on MRI, it is located by inspecting the candidate cartilage structures.

Let (x, y, z) span a coordinate system such that z runs in the AP-direction. The coronal slice that contains the highest number of cartilage voxels is found, by looking for the slice $z = z_s$ that contains the largest translational symmetry in CC-direction. The slice of the binary volume defined by z_s will be referred to as S_{z_s} .

The coordinate x_s is the center of the body in RL-direction.

Finally, the coordinate y_s must define a point on the CC-axis, such that the cartilage structures can be grown from there. We search for the point to which all cartilage structures are oriented. This point will have a large radial symmetry in S_{z_s} . Then y_s will be given by the lowest y -coordinate for which a maximal radial symmetry occurs on the line $x = x_s$, as shown in Fig. 1(d). The radial symmetry transform was implemented as in [9].

In practice, the exact location of the sternum point will not influence the final segmentation as long as it is not too close to another point with large radial symmetry. An offset of 5 voxels still yields similar results.

Preliminary Segmentation. Next, a volume of interest needs to be created around the sternum point (x_s, y_s, z_s) , containing as many cartilage structures as possible. This volume is grown in the y -direction based on S_{z_s} as shown in Fig. 1(e), and is extended in 3D by duplicating it in the surrounding slices. Since the extent of the cartilage structures increases when moving downwards, the volume grows wider every step. The growing stops when the amount of structure found for a larger value of y is below a threshold of 20 voxels. Since the superior part of the cartilage is located more posteriorly, the resulting volume is tilted around the x -axis, to look for the angle at which a maximum overlap with the binary vesselness volume occurs. All connected structures touching this tilted volume are contained in the preliminary segmentation, as shown in Fig. 1(f).

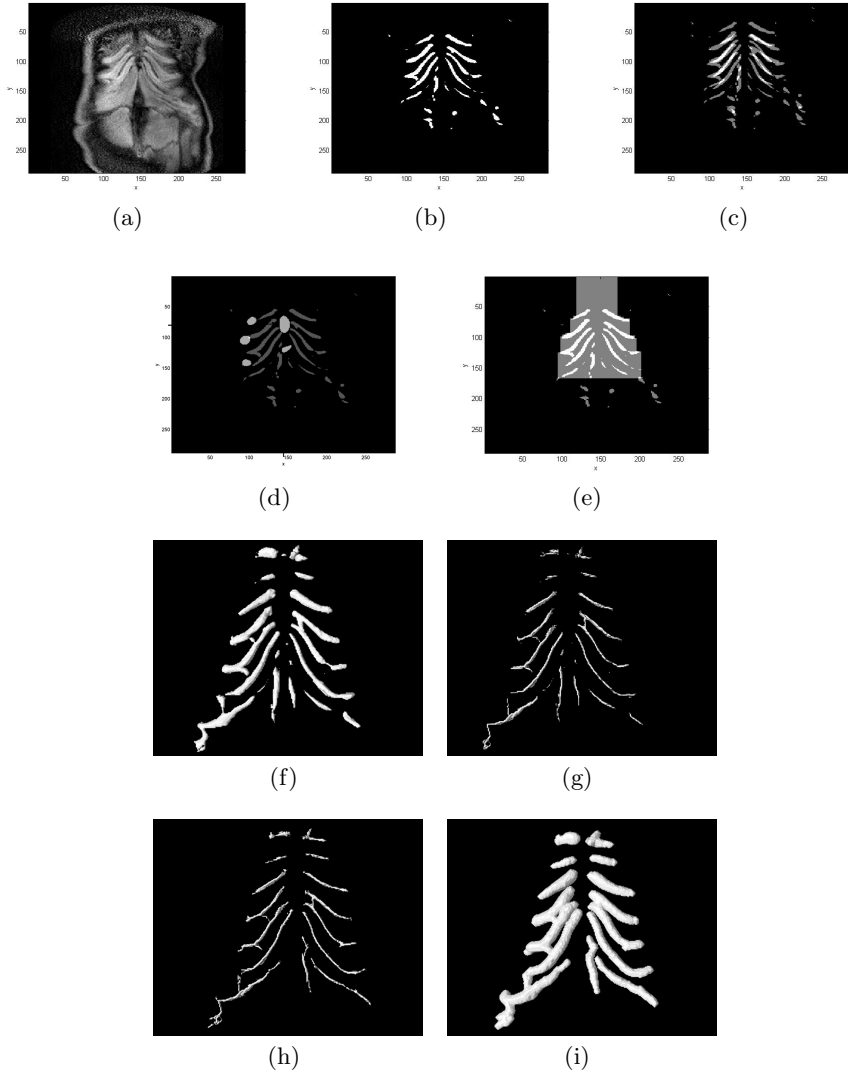


Fig. 1. The different steps of the segmentation pipeline. a) The original volume at slice z_s . b) The slice S_{z_s} is the slice in the thresholded vesselness volume that contains most cartilage structures. c) The symmetry overlap is maximal in this slice. d) The coordinate y_s is given by the first maximum of radial symmetry on the line $x = x_s$. e) The volume of interest is grown from the sternum point on S_{z_s} . f) All connected structures that touch the tilted volume of interest are contained in the preliminary segmentation (3D view). g) The preliminary segmentation is skeletonized (3D view). h) After removing vertical structures and growing line segments together, all cartilage structures are presented by a line element (3D view). i) The final segmentation from the line elements by fast marching level sets (3D view).

Skeletonization and Ridgeness. Ridges are located on the vesselness output, by searching for local ridgeness as in [5]. The preliminary segmentation is skeletonized (Fig. 1(g)). This yields several line structures, some of which belong to the same rib. To be able to grow these segments together, for every voxel in the skeleton, a local neighborhood is investigated to look for ridge voxels that likely belong to the same structure. For seed voxels \mathbf{x}_s and candidate ridge voxels \mathbf{x}_c , with orientations (defined as the unit vector parallel to the ridge) represented by \mathbf{v}_s and \mathbf{v}_c , respectively, the following criteria are used, as in [5]:

1. A candidate voxel should be close to the seed voxel: $\|\mathbf{x}_c - \mathbf{x}_s\| \leq \epsilon_c$.
2. A candidate voxel should have a similar orientation to that of the seed voxel: $\|\mathbf{v}_c \cdot \mathbf{v}_s\| \geq \epsilon_0$.
3. A candidate voxel should not be on a parallel ridge. Therefore, the vector pointing from the seed voxel to the candidate voxel should have a direction similar to the orientation of the candidate voxel: $\|(\mathbf{x}_c - \mathbf{x}_s) \cdot \mathbf{v}_c\| \geq \epsilon_p$.

The thresholds ϵ_c , ϵ_0 and ϵ_p were set to 10, 0.9 and 0.8, respectively.

Level Set Evolution. At this point, all visible cartilage structures in the image will be represented by a line element (Fig. 1(h)). As a final step, level set evolution is performed to yield a full segmentation. For this purpose, the ITK fast marching level set filter is used to dilate the line elements, and finally the ITK Laplacian segmentation level set filter is used to refine the output.

The fast marching level set filter requires two inputs: a set of seed points and a speed function. The seed points are given by the line elements. The speed function is calculated over the anisotropically diffused vesselness output, to prevent leaking to the liver, which has a similar intensity to cartilage. The speed function is obtained by calculating a sigmoid function over all voxels:

$$f(u) = \frac{1}{1 + e^{-\frac{u-\beta}{\alpha}}} , \quad (1)$$

where u is the voxel intensity value and α and β are segmentation parameters. The values of α and β influence the correctness of the speed function. Since we want the speed function to have value 1 inside cartilage structures and 0 on the edges, β should be larger than the value at the edges, and smaller than the value inside. The parameter β is set to 0.2, since this is the average value between the minimum and maximum vesselness output for all images. The parameter α determines the speed at which the function approaches its asymptotes. This value should be small to yield a steep descent in the speed function. We found that $\alpha = \beta/5 = 0.04$ works fine for all cases. The number of iterations was set to 300, which yielded fine results for all cases.

The output of the fast marching level sets serves as an input to the Laplacian segmentation level set filter. Now the anisotropically diffused original image is used as a speed image, on which speed is calculated by Laplacian filtering. This way, the segmentation is allowed to grow towards the cartilage boundaries that

have a low vesselness output. To make sure that it reaches the boundaries, the number of iterations was set to 1000.

The Laplacian level set algorithm yields a final segmentation (Fig. 1(i)).

3 Results

The results are evaluated mainly by visual inspection for completeness and accuracy. For two volunteers, an additional evaluation scan was acquired at a known location, on which the segmentation is overlaid, to measure the boundary error.

The results show that the algorithm was able to detect all cartilage structures that are attached to the sternum and contained in the scan. For the first and fourth volunteer, the first cartilage structure, which is close to the clavicle, was not contained in the field of view and therefore these segmentations counted six structures attached to the sternum, instead of seven. In most scans not all of the 8th, 9th and 10th cartilage structures were visible, either due to low contrast or their position outside the field of view. The structures that were visible were detected. More structures of the lower rib cartilage were visible on the right side near the liver, which is also our main interest.

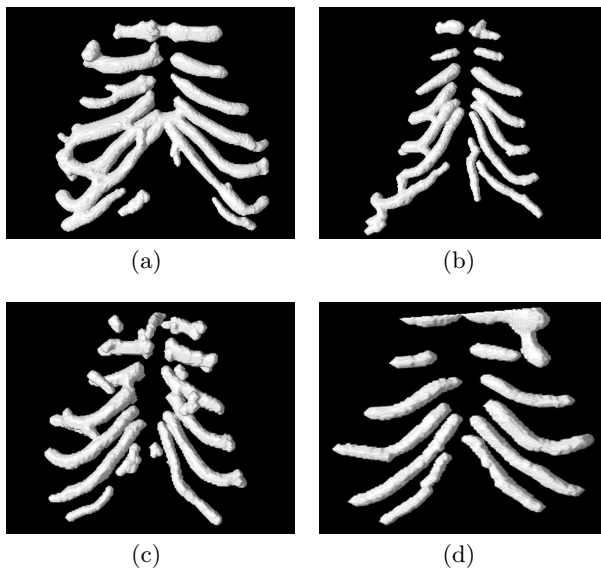


Fig. 2. 3D renderings of the cartilage segmentation for all four volunteers

Results are shown in Fig. 2 and Fig. 3. The 3D rendering is shown for all volunteers. In addition, one slice is taken from every scan, on which the segmentation contours are overlaid, to illustrate the correctness of the segmentation by visual inspection.

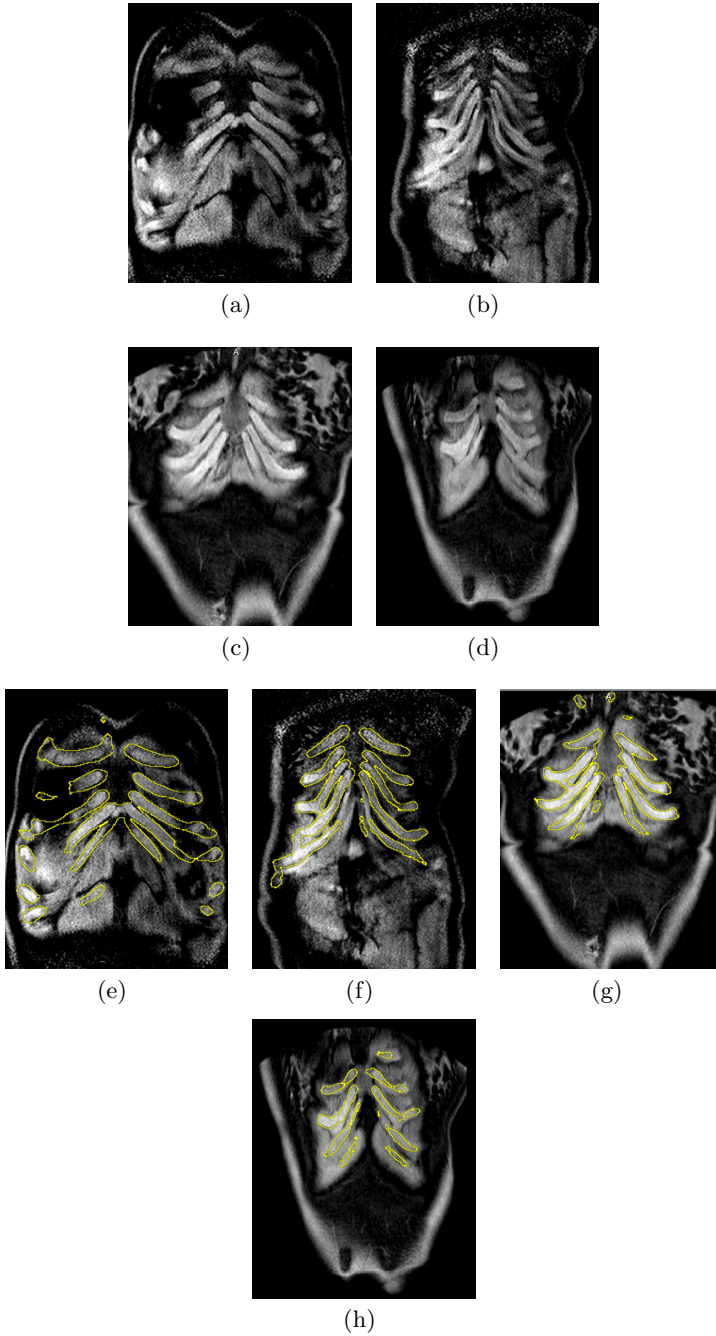


Fig. 3. The evaluation slices for each volunteer (a-d). On these slices, the segmentation contours are overlaid for visual inspection (e-h).

On the third scan, shown in Fig. 3(g), the cartilage of the 6th and 7th left ribs has not been segmented fully. The same holds for the 7th ribs in the second scan (Fig. 3(f)) and the fourth scan (Fig. 3(h)). The vesselness was too low at these locations, due to low contrast. The gaps were too large to be filled by ridge finding, due to the size of the neighborhood used. The size of the neighborhood is limited, since a larger neighborhood will induce more false positives.

In general, the segmentation was accurate. Gaps or false positives may occur, but the algorithm was able to segment almost the entire costal cartilage.

For the third and fourth volunteer, an evaluation scan was acquired at a known location in the body, such that the boundary error of the segmentation can be measured objectively. This scan was angulated such that it is orthogonal to the cartilage that covers the liver, showing the profile of these structures. This way, it is possible to measure the diameter of the structures accurately. The segmentation is overlaid on this scan to evaluate the boundary accuracy, as shown in Fig. 4(a) and Fig. 4(b). The boundaries were segmented manually to serve as a reference standard. The boundary error was measured in AP- and CC-direction on four different ribs, at two locations in each rib, by measuring the distance in voxels from the segmentation boundary to the real cartilage boundary in these directions. The cartilage was oversegmented at some locations, but never undersegmented. For use in the clinic, an undersegmentation could cause the HIFU beam to hit the rib cage. However, an oversegmentation provides an additional safety margin and is in that sense not a problem. The mean boundary error was 1 voxel (resulting in an error of 1.37 mm) in AP-direction and 0.6 voxels (0.82 mm) in CC-direction. The highest outlier was an oversegmentation of 2.5 voxels (3.43 mm). At three locations in both volunteers, the error was 0 mm. The average boundary error was 1.1 mm.

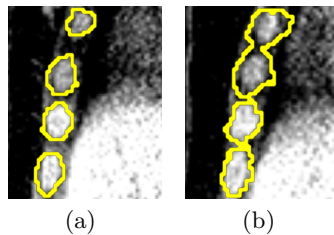


Fig. 4. Evaluation of the boundary accuracy, with contours representing the inside of the segmentation boundary

4 Conclusion

An automatic segmentation algorithm was developed to segment the cartilage structures of the rib cage in MR images. The segmentation scheme is based on analysis of shape and orientation, sternum localization, skeletonization and ridge finding and level set evolution.

The segmentation was evaluated by checking the completeness and the boundary error of the cartilage structures. Since these are preliminary results, the evaluation was done mainly by visual inspection of the contours. The method produced accurate segmentations. The boundary error was evaluated quantitatively for two volunteers, yielding an average error of 1.1 mm.

All segmentations were complete, in the sense that all cartilage structures visible in the MR scan were detected. Gaps may occur when the vesselness output was too low, due to low contrast.

In conclusion, the presented algorithm shows promising results. It is able to detect all visible cartilage structures in MR scans of the thorax. Future work will be dedicated to generating results on a larger data set and providing a detailed quantitative evaluation.

References

1. Tamez-Peña, J.G., Barbu-McInnis, M., Totterman, S.: Knee Cartilage Extraction and Bone-Cartilage Interface Analysis from 3D MRI Data Sets. In: Fitzpatrick, J.M., Sonka, M (eds.) *Proc. SPIE Med. Imaging 2004*, vol. 5370, pp. 1774–1784 (2004)
2. Folkesson, J., Dam, E., Fogh Olsen, O., Pettersen, P.P., Christiansen, C.: Automatic Segmentation of the Articular Cartilage in Knee MRI Using a Hierarchical Multi-class Classification Scheme. In: Duncan, J.S., Gerig, G. (eds.) *MICCAI 2005*. LNCS, vol. 3749, pp. 327–334. Springer, Heidelberg (2005)
3. Zhang, K., Lu, W.: Automatic Human Knee Cartilage Segmentation from Multi-Contrast MR Images Using Extreme Learning Machines and Discriminative Random Fields. In: Suzuki, K., Wang, F., Shen, D., Yan, P. (eds.) *MLMI 2011*. LNCS, vol. 7009, pp. 335–343. Springer, Heidelberg (2011)
4. Yin, Y., Zhang, X., Williams, R., Wu, X., Anderson, D.D., Sonka, M.: LOGISMOS – Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces: Cartilage Segmentation in the Knee Joint. *IEEE T. Med. Imaging* 29, 2023–2037 (2010)
5. Staal, J., van Ginneken, B., Viergever, M.A.: Automatic Rib Segmentation and Labeling in Computed Tomography Scans Using a General Framework for Detection, Recognition and Segmentation of Objects in Volumetric Data. *Med. Image Anal.* 11, 35–46 (2007)
6. Klinder, T., Lorenz, C., von Berg, J., Dries, S.P.M., Bülow, T., Ostermann, J.: Automated Model-Based Rib Cage Segmentation and Labeling in CT Images. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part II*. LNCS, vol. 4792, pp. 195–202. Springer, Heidelberg (2007)
7. Banik, S., Rangayyan, R., Boag, G.: Automatic Segmentation of the Ribs, the Vertebral Column, and the Spinal Canal in Pediatric Computed Tomographic Images. *J. Digit. Imaging* 23, 301–322 (2010)
8. 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)
9. Loy, G., Zelinsky, A.: Fast Radial Symmetry for Detecting Points of Interest. *IEEE Trans. Pattern Anal. Mach. Intel.* 25, 959–973 (2003)

Method for Detecting Enlarged Lymph Nodes from 3D Abdominal CT Images with a Multi-shape and Multi-scale Ellipsoidal Structure Detection Filter

Masahiro Oda¹, Takayuki Kitasaka², Michitaka Fujiwara³,
Kazunari Misawa⁴, and Kensaku Mori^{5,1}

¹ Graduate School of Information Science, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8603, Japan
`moda@is.nagoya-u.ac.jp`

² Faculty of Information Science, Aichi Institute of Technology,
1247 Yachigusa, Yagusa-cho, Toyota, Aichi, 470-0392, Japan

³ Graduate School of Medicine, Nagoya University
65 Tsurumai-cho, Syouwa-ku, Nagoya, Aichi, 466-8550, Japan

⁴ Aichi Cancer Center,

1-1 Kanokoden, Chikusa-ku, Nagoya, Aichi, 464-8681, Japan

⁵ Strategy Office, Information and Communications Headquarters,
Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

Abstract. In this paper, a method for detecting gastric lymph nodes from abdominal CT images is proposed. The positions of metastatic cancers and metastatic lymph nodes should be accurately estimated in order to determine the optimal surgical plan for cancer removal. Ellipsoidal- and spherical-shaped lymph nodes are observed in medical images. However, the detection target of previous lymph node detection methods was only the spherical-shaped lymph nodes. We propose a method for detecting both ellipsoidal- and spherical-shaped lymph nodes by using a multi-shape and multi-scale ellipsoidal structure detection filter that detects the lymph nodes from CT images. The size and the shape of a detection target are specified by the parameters of the filter. The multi-shape and multi-scale detection is performed by applying the filter multiple times with different values for the parameters. Experimental results using 16 cases of CT images showed that the proposed method could detect 56.8% of lymph nodes.

Keywords: CT image, CAD, gastric lymph node detection.

1 Introduction

Cancers develop in organs and can spread to other parts of the body. Cancer cells at the originating site can move through the lymphatic system, the bloodstream, and the abdominal cavity. These cells can then begin to grow in other parts of the body. This process is called metastasis. Metastatic cancers may appear in many

different parts of the body. If a primary cancer is found in diagnosis, physicians need to identify the locations of metastatic cancers to completely cure a patient. Lymph nodes become enlarged if they contain metastatic cancer. The positions and the number of enlarged lymph nodes are examined to clarify the stage of the cancer. Cancers and enlarged lymph nodes around the cancers are removed in surgery to reduce the risk of the cancer recurring [1]. However, surgical removal of excess tissue may decrease the quality of life of a patient. The positions of metastatic cancers and metastatic lymph nodes should be accurately estimated to determine the optimal surgical plan for removing cancer.

To develop a computer aided diagnosis and a surgical planning systems of cancers, an automated detection method of lymph nodes from medical images is required. High resolution 3D medical images such as 3D computed tomography (CT) images are used for cancer diagnoses. Fig. 1 shows examples of lymph nodes in CT images. CT value contrasts between the lymph nodes and the surrounding regions are low. Additionally, organs which have similar CT values to the lymph nodes are usually located close to the lymph nodes. These features make finding the lymph nodes from CT images very difficult.

Methods for detecting lymph node from CT images have been proposed by several research groups. A local intensity structure analysis based on the eigenvalues of the Hessian matrix was used for detecting mediastinal lymph nodes [2] and abdominal lymph nodes [3]. A template matching technique was used for detecting abdominal lymph nodes [4]. Barbu et al. [5] proposed a machine learning and MRF based detection method for axillary lymph nodes. Kitasaka et al. [6] used a 3D minimum directional difference filter to detect lymph nodes in abdominal CT images. These methods detected spherical shaped regions as lymph node candidates in images. The shape of the lymph node is ellipsoidal. The shape changes to spherical if the lymph node enlarges. Thus, ellipsoidal- and spherical-shaped lymph nodes are observed in CT images. However, the detection target of the previous methods was only the spherical-shaped lymph nodes.

In this paper, a method for detecting enlarged gastric lymph nodes from 3D abdominal CT images is proposed. This method detects both ellipsoidal- and spherical-shaped lymph nodes by using a multi-shape and multi-scale ellipsoidal structure detection filter. The lymph nodes are observed as ellipsoidal- or spherical-shaped regions, which have slightly higher CT values than the surrounding regions. We developed the multi-shape and multi-scale ellipsoidal structure detection filter on the basis of local intensity structure analysis in order to detect such lymph nodes. The size and shape of a detection target are specified by the parameters of the filter. Multi-shape and multi-scale detection is performed by applying the filter multiple times with different values for the parameters.

2 Method

2.1 Overview

The proposed method consists of preprocessing, detection, and false positive (FP) reduction processes. In the preprocessing step, we extract regions that

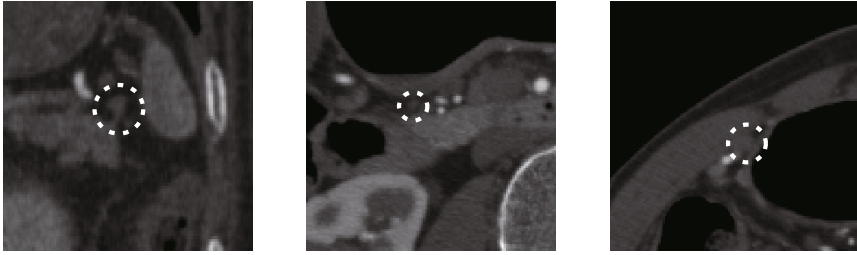


Fig. 1. Examples of lymph nodes in CT images. The lymph nodes are indicated by circles.

define which of the following processes are applied. The multi-shape and multi-scale ellipsoidal structure detection filter is used in the detection process. Final detection results are obtained after the FP reduction process.

2.2 Preprocessing

The input of the proposed method is a 3D contrasted abdominal CT image of a patient who has gastric cancer. We apply a 3D Gaussian smoothing filter to the CT image. The standard deviation of the Gaussian smoothing filter is 0.5 [mm]. The detection process is applied to the soft tissue of the body. We extract a soft tissue region from the CT image by using a thresholding process. The soft tissue region R is extracted as a set of voxels having CT values from -400 to 200 [H.U.] in the CT image.

2.3 Detection Process

Enlarged lymph nodes have higher CT values than the surrounding voxels. The shape of enlarged lymph nodes is similar to an ellipsoid or sphere. We detect enlarged lymph nodes having such features by using the multi-shape and multi-scale ellipsoidal structure detection filter.

We approximate the local CT value distribution with a quadratic polynomial of three variables by fitting it to the neighborhood (cubic region whose edge length is l [mm]) of each voxel in R . The surface fitting is performed by minimizing the fitting error, which is defined as the squared difference of the approximated and actual CT values. We obtain a Hessian matrix consisting of second-order partial differential coefficients of the polynomial. Let the eigenvalues of the Hessian matrix be λ_1 , λ_2 , and λ_3 ($\lambda_3 \leq \lambda_2 \leq \lambda_1$) and their corresponding eigenvectors be \mathbf{e}_1 , \mathbf{e}_2 , and \mathbf{e}_3 . Eigenvalues λ_1 and λ_3 correspond to eigenvectors \mathbf{e}_1 and \mathbf{e}_3 , along which the second derivatives are maximum and minimum. λ_2 is a second derivative value of a direction orthogonal to eigenvectors \mathbf{e}_1 and \mathbf{e}_3 . When the distribution of local CT values around a target voxel shows an ellipsoidal structure, the eigenvalues of the Hessian matrix show the pattern $\lambda_3 < \lambda_2 \approx \lambda_1 < 0$. Fig. 2 shows profiles of eigenvalues calculated on

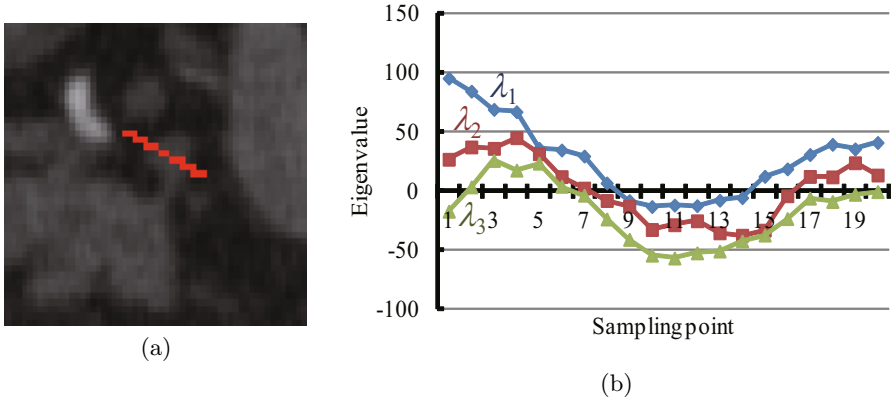


Fig. 2. (a) An example of a lymph node. (b) Profiles of eigenvalues calculated at each voxel on the red line shown in the figure (a). Eigenvalues in the lymph node are shown near the center of the graph. The eigenvalues show the pattern in the ellipsoidal structure ($\lambda_3 < \lambda_2 \approx \lambda_1 < 0$).

an actual CT image. We define the ellipsoidal structure detection filter, which detects voxels in the ellipsoidal structure showing the condition of eigenvalues, as

$$S(\lambda_1, \lambda_2, \lambda_3) = |\lambda_3| \left(1 - \frac{1}{1-p} \left| \frac{\lambda_2}{\lambda_3} - p \right| \right)^\gamma \psi(\lambda_1, \lambda_2), \quad (1)$$

where

$$\psi(\lambda_s, \lambda_t) = \begin{cases} \left(\frac{\lambda_s}{\lambda_t} \right)^\gamma, & \text{if } \lambda_t \leq \lambda_s < 0, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Equation 2 was proposed by Sato et al. 7. The parameter p defines the numerical relation of λ_2 and λ_3 ($0 \leq p < 1$). The absolute values of λ_2 and λ_3 represent the magnitude of CT value changes along the long and short axes of the ellipsoidal structure, respectively. The positional relationship of an ellipsoidal structure and the eigenvectors is shown in Fig. 3. The parameter γ defines the sharpness of selectivity for the ellipsoidal structure.

The structure of a detection target of the filter is changed according to the parameter p . The filter detects ellipsoidal structures that have a shape similar to a sheet structure if the parameter p is close to 0. The filter also detects ellipsoidal structures that have a shape similar to a spherical structure if the parameter p is close to 1. The scale of the detection target of the filter is specified by the parameter l . The filter is applicable to detecting various shapes and scales of detection targets by selecting the proper values of p and l .

We apply the ellipsoidal structure detection filter to each voxel in the region R . We generate a detection result image I_d (a grayscale image with background voxels) that has output values of the filter on each voxel. Lymph node candidate regions are extracted as a set of voxels whose values are greater than 0 in I_d .

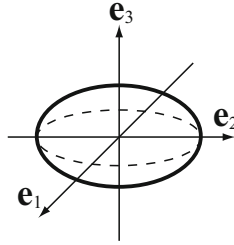


Fig. 3. Directions of the eigenvectors in an ellipsoidal structure. \mathbf{e}_1 and \mathbf{e}_2 correspond to the long axis, and \mathbf{e}_3 corresponds to the short axis of the ellipsoidal structure.

We calculate the output values of the filter with different settings of parameters p and l for a voxel in order to perform a multi-shape and multi-scale detection. The maximum output value of the filter within the output values of the filter calculated with different settings of the parameters is stored in the I_d .

2.4 FP Reduction Process

To reduce FPs in the detection result stored in I_d , we perform thresholding processes based on voxel values, volumes, maximum voxel values, distances from stomach regions, and contrasts of CT values between a candidate region and the surrounding region.

All voxels whose intensity value is lower than t_1 are removed. The rest of the processes are region-based reduction processes. Small connected components whose volume is lower than t_2 [mm³] and connected components where the maximum intensity value is lower than t_3 are removed. Connected components more than t_4 [mm] away from the stomach region are removed. We extract the stomach region manually. Automation of this extraction process is easy if the inside of the stomach is filled with air. However, the stomach may contain fluid or food, which causes automated extractions to fail. Therefore, we perform the extraction manually.

The difference in average CT values between the lymph node candidate region and the surrounding region is also used to remove FPs. We calculate the average CT value a_1 [H.U.] of voxels in a candidate region. A dilation filter, whose structure element is a sphere 3.0 [mm] in radius, is applied to the candidate region. We obtain the surrounding region of the candidate region by subtracting the candidate region from the result of the dilation filter. The average CT value of voxels in the surrounding region is described as a_2 [H.U.]. The candidate region is removed if the average CT values satisfy either conditions $a_1 - a_2 < t_5$ or $t_6 < a_1 - a_2$.

The remaining regions after these reduction processes are the final lymph node candidate regions.

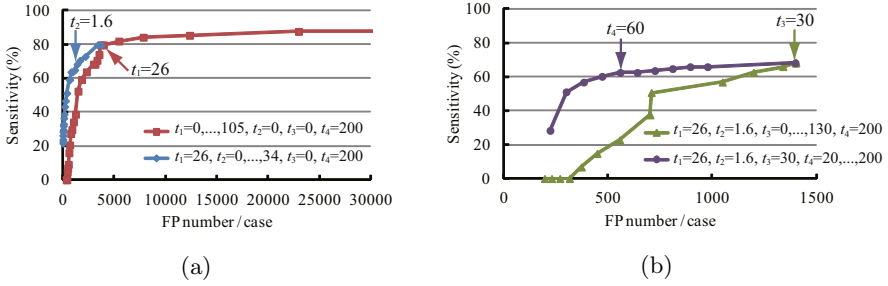


Fig. 4. FROC curves with parameters: (a) (red curve) $t_1 = 0, \dots, 105, t_2 = 0.0, t_3 = 0,$ and $t_4 = 200$; (a) (blue curve) $t_1 = 26, t_2 = 0.0, \dots, 34.0, t_3 = 0,$ and $t_4 = 200$; (b) (green curve) $t_1 = 26, t_2 = 1.6, t_3 = 0, \dots, 130,$ and $t_4 = 200$; and (b) (purple curve) $t_1 = 26, t_2 = 1.6, t_3 = 30,$ and $t_4 = 20, \dots, 200.$

3 Experiments

We applied the proposed method to 16 cases of contrasted 3D abdominal CT images. All were stomach cancer cases. 88 lymph nodes were found by a medical doctor in the CT images. The acquisition parameters of the CT images were image size: 512×512 [pixels], number of slices: 359-584, pixel spacing: 0.625-0.741 [mm], slice spacing: 0.500-0.800 [mm], slice thickness: 0.500-1.00 [mm], tube voltage: 120 [kVp], and tube current: 150-450 [mAs]. We experimentally determined the parameters of the method as $l = \{4.0, 5.0, 6.0\}$ [mm], $p = \{0.7, 0.9\}$, and $\gamma = 1.$

To determine the parameters of the FP reduction process, we generated FROC curves by calculating the sensitivity and the FP number/case of various FP reduction parameters. Fig. 4 (a) shows a FROC curve (red color) obtained by setting the parameters as $t_1 = 0, \dots, 105, t_2 = 0.0, t_3 = 0,$ and $t_4 = 200.$ From the FROC curve, sensitivity was 79.5% with 3739.9 FPs/case when $t_1 = 26.$ Similarly, from the FROC curve shown in Fig. 4 (a) (blue color), sensitivity was 68.2% with 1401.1 FPs/case when $t_2 = 1.6$ [mm³]. From the FROC curve shown in Fig. 4 (b) (green color), sensitivity was 68.2% with 1398.3 FPs/case when $t_3 = 30.$ From the FROC curve shown in Fig. 4 (b) (purple color), sensitivity was 62.5% with 558.6 FPs/case when $t_4 = 60$ [mm]. t_5 and t_6 were also determined from FROC curves as $t_5 = 50$ and $t_6 = 110.$ In summary, the sensitivity of the proposed method was 56.8% (50/88) with 262.3 FPs/case.

Examples of lymph nodes detected by the proposed method are shown in Fig. 5

4 Discussion

The proposed method detected lymph nodes from contrasted 3D abdominal CT images. The ellipsoidal structure detection filter could detect the various shapes and scales of the lymph nodes by setting various values to parameters p and

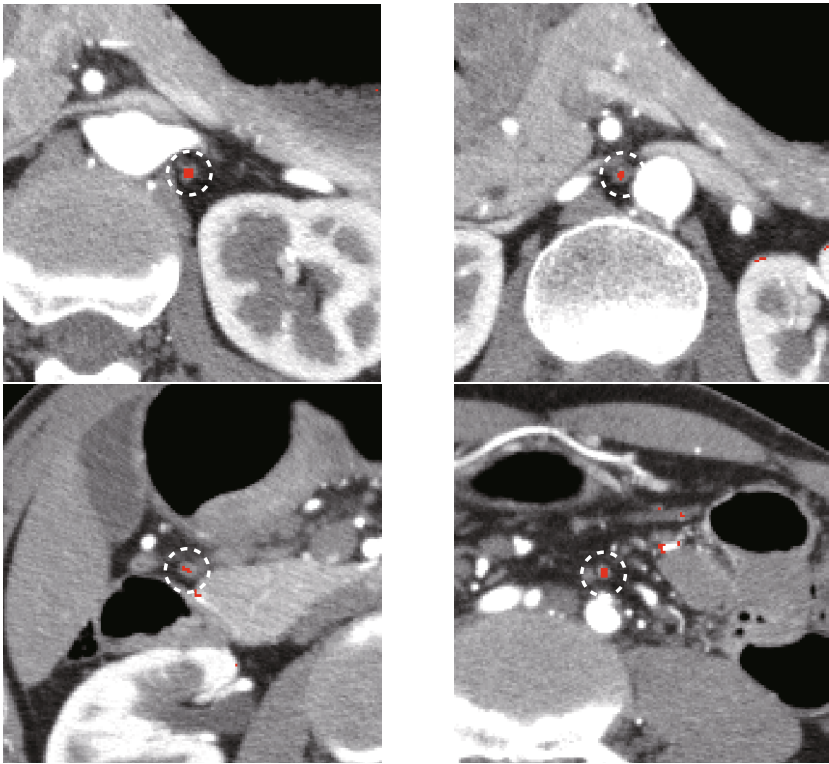


Fig. 5. Detection results of the proposed method. The size of the detection target was from 4.0 to 6.0 [mm]. The marked regions are detected regions. The regions indicated by the circles are true positive regions.

l. The sensitivity of the abdominal lymph node detection method proposed by Mori et al. [3] was 79.2% with 98 FPs/case, and that of the axillary lymph node detection method proposed by Barbu et al. [5] was 82.3% with 1 FP/case. However, the sizes of their detection targets were 5.0 [mm] or larger [3] and 10.0 [mm] or larger [5]. The size (scale) of our detection target was from 4.0 to 6.0 [mm]. Our method can detect very small lymph nodes in comparison to the previous methods. The sensitivity of our method (56.8%) is close to that of humans finding lymph nodes from CT images. Our method has the potential to help in the diagnosis of metastatic cancers.

Some of the enlarged lymph nodes had ellipsoidal shapes. However, most of the enlarged lymph nodes had spherical shapes in the CT images. We set the values {0.7, 0.9} to the parameter p to detect spherically shaped lymph nodes. The scale of the detection target was 4.0 to 6.0 [mm]. Though these parameters were determined experimentally, they should be determined on the basis of the actual shapes and scales of the lymph nodes. To obtain the optimal values of parameters p and l , the statistical analysis of shapes and scales of lymph nodes using a larger number of CT images is necessary.

Although we performed the FP reduction process on the basis of the thresholding processes of the output values of the filter and CT values, the detection results still contained many FPs. To reduce FPs, it is necessary to introduce to the FP reduction process new feature values such as the shape of a candidate region and the positional relationship of candidate regions. Using the anatomical information of the lymph nodes, such as the average position of the lymph node and the positional relationships between the lymph nodes and other organs, in the FP reduction process will be effective for further reducing FPs. Such anatomical information based detection requires the statistical analysis of lymph nodes and organ extractions.

5 Conclusion

In this paper, a method for detecting gastric lymph nodes from abdominal CT images was proposed. We introduced a multi-shape and multi-scale ellipsoidal structure detection filter that detects lymph nodes. The filter can detect the various scales of ellipsoidal and spherical lymph nodes. Experimental results using 16 cases of CT images showed that the proposed method can detect 56.8% of lymph nodes. Future work includes the development of a FP reduction process based on anatomical information, optimization of the parameters, and evaluation of the method by using a larger number of images.

References

1. Ozaki, M.: Preoperative diagnosis of lymphmetastasis of gastric cancer using CT scan. *The Japanese Journal of Gastroenterological Surgery* 17(8), 1507–1516 (1984)
2. 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: *Proc. SPIE Medical Imaging*, vol. 7260, 72600V-1–72600V-11 (2009)
3. Mori, K., Nakamura, Y., Kitasaka, T., Ito, M., Nawano, S., Misawa, K., Fujiwara, M.: Automated Detection And Display of Abdominal Lymph Nodes From CT Volumes Based on Local Intensity Structure Analysis. *International Journal of Computer Assisted Radiology and Surgery* 5(1), S125–S126 (2010)
4. Masakiyo, F., Kubo, M., Kawata, Y., Niki, N., Nishioka, M., Miyamoto, H., Yoshikawa, K., Shimada, M., Nishitani, H.: Extraction Algorithm of Enlargement Lymph Nodes Using Abdominal CT Datasets. In: *IEICE Technical Report*, MI2008-10, pp. 55–58 (2008)
5. Barbu, A., Suehling, M., Xu, X., Liu, D., Zhou, S.K., Comaniciu, D.: Automatic Detection and Segmentation of Axillary Lymph Nodes. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) *MICCAI 2010, Part I. LNCS*, vol. 6361, pp. 28–36. Springer, Heidelberg (2010)
6. 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, pp. 336–343. Springer, Heidelberg (2007)
7. Sato, Y., Westin, C.F., Bhalerao, A., Nakajima, S., Shiraga, N., Tamura, S., Kikins, R.: Tissue Classification Based on 3D Local Intensity Structures for Volume Rendering. *IEEE Transactions on Visualization and Computer Graphics* 6(2), 160–180 (2000)

Delineation of Liver Tumors from CT Scans Using Spectral Clustering with Out-of-Sample Extension and Multi-windowing

Jiayin Zhou¹, Weimin Huang¹, Wei Xiong¹, Wenyu Chen¹,
Sudhakar K. Venkatesh², and Qi Tian¹

¹ Institute for Infocomm Research, Agency for Science, Technology and Research,
1 Fusionopolis Way, #21-01 Connexis South Tower, Singapore 138632, Singapore
jzhou@i2r.a-star.edu.sg

² Department of Radiology, Mayo Clinic,
200 First St. S.W., Rochester, Minnesota 55905, USA

Abstract. Accurate extraction of live tumors from CT data is important for disease management. In this study, an algorithm based on spectral clustering with out-of-sample extension is developed for the semi-automated delineation of liver tumors from 3D CT scans. In this method, spatial information is incorporated into a similarity metric together with low-level image features. A trick of out-of-sample extension is performed to reduce the computational burden in eigen-decomposition for a large matrix. Experimental results show that the developed method using multi-windowing feature obtained better results than using only the original data-depth and the support vector machine method, with a sensitivity of 0.77 and a Jaccard similarity measure of 0.70.

Keywords: Spectral clustering, out-of-sample extension, CT, tumor delineation.

1 Introduction

Liver tumors include benign tumors, primary malignant tumors and liver metastases. Contrast-enhanced multi-detector computed tomography (MDCT) is widely used for the detection, diagnosis and management of liver tumor. Accurate delineation of liver tumor in 3D CT scans is important for tumor volume quantification, which is used for disease prognosis and assessment of treatment response. Manual tumor margin delineation is very tedious, especially for thin-section CT scans with a number of tumors in one patient. Some semi-automated methods, including region growing, watershed, level-set, hidden Markov measure field, etc, have been developed to segment liver tumor from CT data [1-5]. The majority of these methods segment each tumor individually in 2D slices or 3D volumes, hence the efficiency is not high when processing scans with multiple lesions. It is desired to develop new techniques which are able to extract and delineate all possible livers tumor simultaneously from a 3D CT data volume. This will greatly assist clinicians in liver tumor diagnosis, treatment planning and follow-up management.

In general, for the standard 4-phase contrast-enhanced liver CT scan, images from the portal venous (PV) phase give the best visual discrimination for liver objects including liver parenchyma, liver vessels (hyperdense) and tumors (hypodense, if has). Hence it is interesting to explore clustering-based method to classify or group voxels into different tissue classes according to certain similarity criteria. The applications of spectral clustering approach [6, 7] in the segmentations of magnetic resonance and positron emission tomography images have been reported [8-10]. The main idea of spectral clustering is to find the principal eigenvectors and eigenvalues, i.e. the spectrum, of a large affinity matrix encoding the pairwise similarity of voxels and then to group these principal eigenvectors to achieve the graph partitioning. This equivalently accomplishes the grouping of voxels based on their pairwise similarity and the paths they form in the graph represented by the affinity matrix.

To the best of our knowledge, there is no report on spectral clustering method for the segmentation of CT data. In this paper, we present a spectral clustering-based method to delineate liver tumors from CT volumes in PV phase at a semi-supervised manner, with the following three emphases. First, a pairwise similarity metric combining both spatial and low-level image features is proposed for the spectral clustering framework. Secondly, we incorporate out-of-sample extension [11] into spectral clustering, hence the main computational burden of eigen-decomposition of the affinity matrix can be avoided and the segmentation can be achieved at the original image resolution, leading to more accurately segmented regions. Thirdly, original CT data volume in 12-bit data depth is re-scaled by multi-windowing, to obtain a set of images with different contrasts as the low-level feature for clustering work. Our proposed method was tested by 12 MDCT data volumes and the results were quantitatively evaluated.

2 Method

2.1 Framework of Spectral Clustering

Spectral clustering is often considered as an approximate solution to graph-cut problem that obtains data partitioning by cutting weak links between graph nodes to separate data into various segments. The framework of the algorithm is as follows [6]:

Given a set of data points $S = \{s_1, \dots, s_n\}$ in R^d to be clustered into k subsets,

- (1) Form the affinity matrix defined by $A_{ij} = \exp(-\|s_i - s_j\|^2 / 2\sigma^2)$ if $i \neq j$, and $A_{ii} = 0$;
- (2) Define D to be a diagonal matrix whose (i, i) -element is the sum of A 's i -th row, and construct the matrix $L = D^{-1/2} A D^{-1/2}$;
- (3) Find x_1, x_2, \dots, x_k , the k largest eigenvectors of L , and form the matrix $X = [x_1 \ x_2 \ \dots \ x_k] \in R^{n \times k}$, by stacking the eigenvectors in columns;
- (4) Renormalize each of X 's rows to have the unit length to form the matrix Y from X ;
- (5) Treating each row of Y as a point in R^k , the rows in Y are clustered into k clusters via the K -means algorithm;

- (6) Assign the original points s_i to cluster j only if row i of the matrix Y was assigned to cluster j .

In most of the cases, the affinity matrix is formed by the Euclidean distances between pairwise data points, scaled by $2\sigma^2$, to reflect the degree of similarity between pairwise data points in S .

2.2 A Similarity Metric Combining Spatial and Low-Level Features

Being independent of any assumption of a distributional model, the success of a spectral clustering method depends heavily on the choice of an appropriately selected metric that measures the pairwise similarity between data points. In this study, we construct a metric $d(s_i, s_j)$ that reflects both the spatial and low-level features between two voxel s_i and s_j in a CT volume:

$$\begin{aligned}
 d(s_i, s_j) &= K(s_i, s_j) \cdot G(s_i, s_j), \\
 d(s_i, s_j) &= \exp\left(-\alpha \|p_i - p_j\|^2 / 2\sigma_1^2\right) \cdot \exp\left(-\beta \|q_i - q_j\|^2 / 2\sigma_2^2\right), \\
 d(s_i, s_j) &= \exp\left(-\alpha \|p_i - p_j\|^2 / 2\sigma_1^2 - \beta \|q_i - q_j\|^2 / 2\sigma_2^2\right), \quad (1)
 \end{aligned}$$

where σ_1 and σ_2 are scalars, p_i is the spatial location of voxel s_i , q_i is the feature vector based on low-level image features of s_i , α and β are weights fulfilling $\alpha, \beta \in [0, 1]$ and $\alpha + \beta = 1$. Besides of the low-level image features, this metric considers the spatial constraints that tumor is generally a continuous mass in 3D space. Hence given similar low-level image features, two spatially nearby voxels will have more influence on the similarity measure than two voxels that are far apart. Here α and β are used to weight the influences from spatial and low-level image features.

2.3 Out-of-Sample Extension

The main computational burden in spectral clustering approach is the eigen-decomposition of affinity matrix of size $n \times n$, where n equals to the number of voxels in the volume data and the computation complexity is in the order of n^3 . If directly applying spectral clustering to group all the voxels in a CT volume, the dimension of the affinity matrix can be extremely high, causing computational problems in eigen-decomposition. With out-of-sample extension method [11], we only need to find the spectral embedding of a subset of the data voxels from sample selection, while the remaining voxels will be mapped to the spectral space using a mapping function and clustered unsupervisedly therein. Hence, the main computational burden of eigen-decomposition of the similarity matrix can be avoided and the data partitioning can be achieved at the original data resolution. In addition by this step, the unsupervised spectral clustering is converted to a semi-supervised procedure where sample selection

to form the training set is required. The framework of out-of-sample extension for spectral clustering is as follows:

Assume that a set of labeled samples Tr was chosen from the CT volume. Spectral clustering described in section 2.1 was applied into Tr using the similarity metric defined in equation (1), and X and V are the matrices of eigenvector and eigenvalue obtained on Tr . For the remaining data points Te , a “mutual” affinity matrix $D(i, j)$ can be constructed by the similarity metric in equation (1) between pairwise Te and Tr , where $D(i, j)$ is of $m \times l$ size, m and l are the number of data points in Te and Tr , respectively. $D(i, j)$ can be normalized by

$$D(i, j) = \frac{D(i, j)}{\sqrt{\sum_{x \in Tr} D(i, x) \sum_{x' \in Tr} D(j, x')}}. \quad (2)$$

The spectral embedding for Te can be computed from

$$X_{mapped} = D \cdot X \cdot V^{-1}. \quad (3)$$

After obtaining X_{mapped} , clustering algorithms such as K -means can be used to partition the data. Thus spectral clustering can be achieved by out-of-sample extension technique without eigen-decomposition of the very large affinity matrix.

2.4 Rescale CT Data by Multi-windowing

For CT data analysis, voxel density (indicated by Hounsfield Unit, HU) and its derivatives are the most common low-level image features. The voxel density is of 12-bit integer in data depth, but for display and interpretation purposes, CT data are converted into images of 8-bit integer in data depth, with different window level and width settings (*window level, window width*) to highlight different anatomic structures. For example, (-600, 1600) is the standard lung window by which the lung structure, vessels and bronchus can be viewed clearly; (50, 350) is the standard abdominal window by which the organs/structures in abdominal region can be well visualized. To examine liver region, especially to detect liver lesions, a window width of 350 is still too wide so that some tumors may be overlooked due to the close image intensities shown in this window. In practice, radiologists often use a narrower window width to observe liver region. Here we adopt a customized ‘liver window’ [12] for CT data rescaling: Liver window has a window center equal to the density level of liver parenchyma (obtained by online sample selection) and a window width of 180.

3 Implementation and Experiment

Our proposed method was evaluated by using 12 MDCT data volumes (image size of 512×512 pixels, slice thickness of 2-3 mm) acquired in PV phase. The reference standards (RS) of liver tumors were identified and manually traced out by an experienced abdominal radiologist.

Before using the developed method to delineate liver tumors, an anisotropic diffusion filter was applied to CT data to reduce CT image noise but still preserve the significant parts of the image content. Then a 3D flipping-free mesh deformation algorithm [10] was employed for the segmentation of the gross liver as the region-of-interest (ROI). Within the segmented liver ROI, approximately 3 tissue classes exist: normal liver parenchyma, enhanced vessel and tumor, based on prior knowledge. For each data set, samples of each class (tumor, vessel and normal liver parenchyma) were obtained by mouse-pen selection, as shown in Fig. 1. Each 12-bit CT data volume was rescaled into 2 set of 8-bit images using standard abdominal window (50, 350) and the customized ‘liver window’. The corresponding image intensities from the 2 sets of images form a concurrent feature vector (I_{AW}, I_{LW}) to be used as the low-level image feature. In the implementation, both spatial and low-level image features were normalized into [0, 1]. In addition, α and β were set as 0.2 and 0.8, respectively. For comparison purpose, CT data were also tested by 1) the proposed method but using the original data depth, and 2) a binary support vector machine (SVM) method using multi-windowing feature. For the binary SVM method, tumor samples were used as the positive samples and vessel and liver parenchyma samples were combined together as the negative samples.

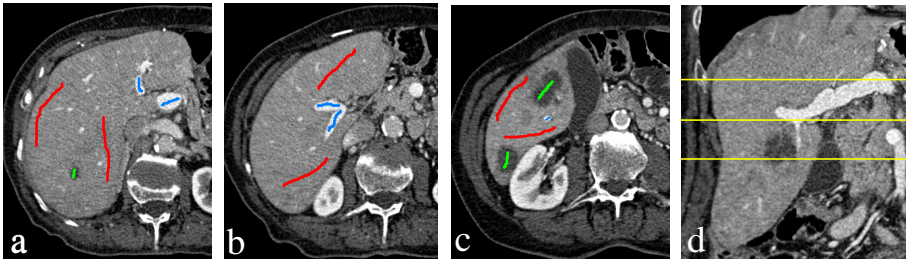


Fig. 1. (a)-(c) Sample selection was performed on 3 slices from different sections, red for normal liver parenchyma, green for lesion, blue for vessel; (d) the sections of the 3 slices in the coronal view, (a) to (c) is from top to bottom.

The extracted lesions were compared with the RS by voxel overlapping on the basis of per data volume (not per tumor). Two quantitative measures, sensitivity (SEN) and Jaccard similarity measure (JSM) were calculated to assess the similarity between the computerized and manually defined liver tumors:

$$SEN = (Vol_{Alg} \cap Vol_{RS}) / Vol_{RS}, \tag{4}$$

$$JSM = (Vol_{Alg} \cap Vol_{RS}) / (Vol_{Alg} \cup Vol_{RS}), \tag{5}$$

where Vol_{Alg} is the tumor volume delineated by the algorithm and Vol_{RS} is the RS . Compared to SEN , JSM considers the impact of false positives on the overall performance of the methods.

4 Results

Tumor delineation was tested on 12 CT data volumes using the proposed method and the SVM method. Fig. 2 shows 3 CT slices in two different windows and the corresponding results after spectral clustering using multi-windowing feature and the original data depth. Fig. 3 shows 3 examples of tumor delineation results using the developed method (with/without multi-windowing feature) and the SVM method. Compared with the other two methods, the developed method with multi-windowing feature is more sensitive in capturing small and peripheral lesions which are with weak image appearance and features, as marked by yellow circles in Fig. 3. In fact, some tiny vessel branches with weak image appearance are also identified by the proposed method. On the contrary they are classified as the liver parenchyma, the class with the majority of elements, by the other methods. One needs to pay attention that in some cases, due to the timing in the acquisition of PV phase images, part of the inferior vena cava (IVC) embedded in the liver is not well enhanced, therefore this part tends to be classified into tumor class by these methods, as shown in the bottom row of Fig. 3. All the three methods misclassified the un-enhanced IVC, which is

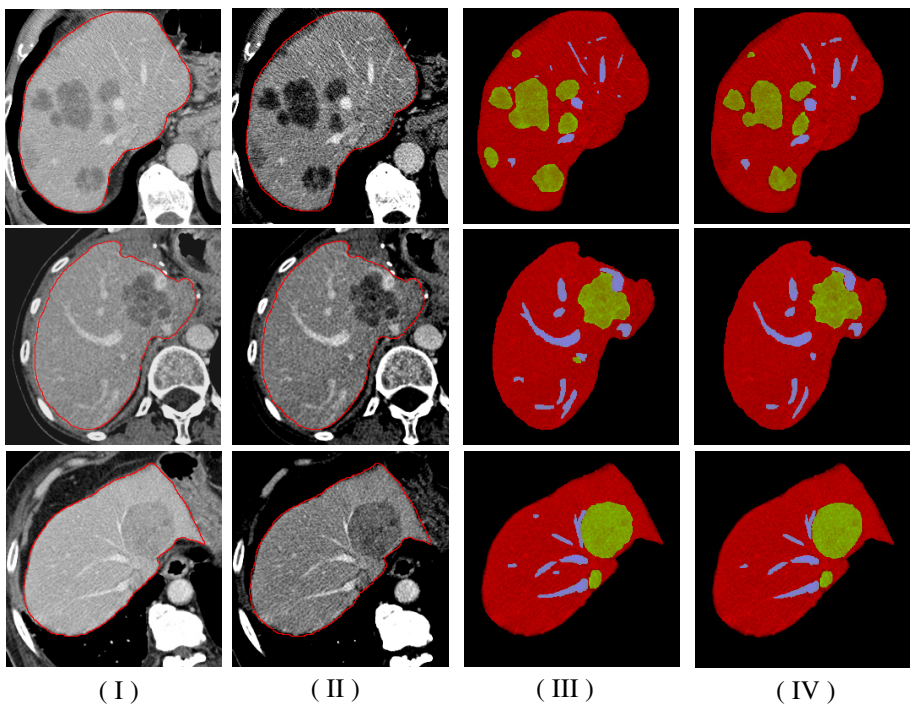


Fig. 2. Three CT slices in abdominal window (I), liver window (II), and the corresponding results after spectral clustering using multi-windowing feature (III) and the original data-depth (IV); red: liver parenchyma; green: tumor, blue: vessel

Table 1. Results of the evaluation metrics expressed as *SEN*, and *JSM*

Data No.	Proposed method with multi-windowing		Proposed method with 12-bit data depth		SVM method	
	<i>SEN</i>	<i>JSM</i>	<i>SEN</i>	<i>JSM</i>	<i>SEN</i>	<i>JSM</i>
1	0.77	0.72	0.80	0.75	0.76	0.72
2	0.62	0.56	0.61	0.56	0.60	0.58
3	0.71	0.61	0.62	0.59	0.70	0.64
4	0.76	0.67	0.57	0.55	0.61	0.55
5	0.79	0.67	0.75	0.68	0.73	0.71
6	0.68	0.63	0.67	0.62	0.68	0.64
7	0.84	0.76	0.73	0.70	0.71	0.70
8	0.74	0.70	0.69	0.59	0.71	0.63
9	0.86	0.83	0.78	0.75	0.82	0.70
10	0.80	0.77	0.71	0.65	0.73	0.69
11	0.78	0.74	0.72	0.69	0.69	0.66
12	0.83	0.79	0.80	0.76	0.82	0.75
Mean±SD	0.77±0.07	0.70±0.08	0.70±0.08	0.66±0.08	0.71±0.07	0.66±0.06

marked by the green arrow in (I), into the tumor class, as marked by green circles. The mean values of *SEN* and *JSM* obtained by the developed method with multi-windowing feature are 0.77 and 0.70, respectively, whereas they are 0.70 and 0.66 for the proposed method without multi-windowing feature, and 0.71 and 0.66 for the SVM method, as shown in Table 1. These figures indicate that the proposed method outperform the other two methods moderately.

5 Conclusion

In this paper, a semi-supervised spectral clustering algorithm with out-of-sample extension is proposed for the delineation of liver tumors from 3D CT data. In this method, a metric combining both spatial information and low-level image features was constructed to encode the pairwise similarity of each data point. A trick of out-of-extension was employed for the reduction of computation burden in matrix eigen-decomposition. In addition, CT data was rescaled into different contrasts by multi-windowing to form the low-level image feature for better discrimination. Experiments were conducted on 12 CT scans and the results suggested that the proposed method with multi-windowing feature performed modestly better than the same method without multi-windowing feature and the SVM method. The improvement is mainly in the increase of detection sensitivity for small and peripheral lesions.

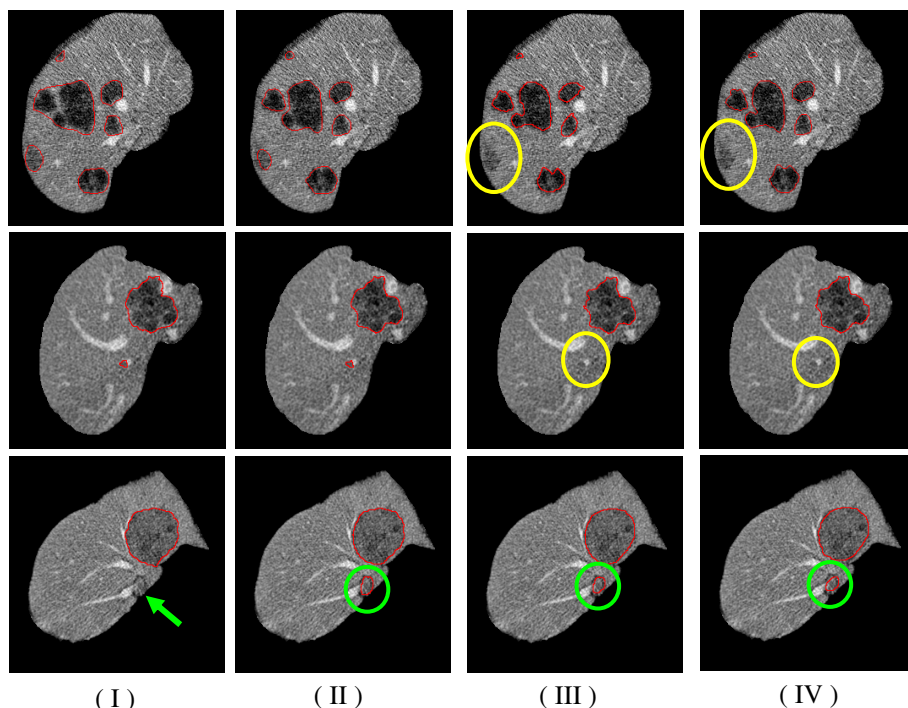


Fig. 3. Three examples of tumor delineation results: (I) tumor *RS*; (II) results using the proposed method with multi-windowing feature; (III) results using the proposed method with the original data-depth; (IV) results using SVM

Acknowledgements. This work was supported by a research grant (JCOAG03_FG05-2009) from the Joint Council Office, Agency for Science, Technology and Research, Singapore.

References

1. Zhao, B., Schwartz, L.H., Jiang, L., et al.: Shape-constraint region growing for delineation of hepatic metastases on contrast-enhanced computed tomograph scans. *Invest. Radiol.* 41, 753–762 (2006)
2. Ray, S., Hagge, R., Gillen, M., et al.: Comparison of two-dimensional and three-dimensional iterative watershed segmentation methods in hepatic tumor volumetrics. *Med. Phys.* 35, 5869–5881 (2008)
3. Smeets, D., Loeckx, D., Stijnen, B., et al.: Semi-automatic level set segmentation of liver tumors combining a spiral-scanning technique with supervised fuzzy pixel classification. *Med. Image Anal.* 14, 13–20 (2010)
4. Hame, Y., Pollari, M.: Semi-automatic liver tumor segmentation with hidden Markov measure field model and non-parametric distribution estimation. *Med. Image Anal.* 16, 140–149 (2012)

5. Li, B.N., Chui, C.K., Chang, S.K., Ong, S.H.: A new unified level set method for semi-automatic liver tumor segmentation on contrast-enhanced CT images. *Expert Syst. Appl.* 39, 9661–9668 (2012)
6. Ng, A.Y., Jordan, M.I., Weiss, Y.: On spectral clustering: analysis and an algorithm. In: *Proc. NIPS*, pp. 849–856 (2001)
7. von Luxburg, U.: A tutorial on spectral clustering. *Stat. Comput.* 17, 395–416 (2007)
8. Carballido-Gamio, J., Belongie, S.J., Majumdar, S.: Normalized cuts in 3D for spinal MRI segmentation. *IEEE Trans. Med. Imaging* 23, 36–44 (2004)
9. Crum, W.R.: Spectral clustering and label fusion for 3D tissue classification: sensitivity and consistency analysis. In: *Ann. BMVA 2009*, pp. 1–12 (2009)
10. Yang, F., Grigsby, P.W.: Delineation of FDG-PET tumors from heterogeneous background using spectral clustering. *Eur. J. Radiol.* (in press)
11. Bengio, Y., Paiement, J., Vincent, P.: Out-of-sample extensions for LLE, isomap, MDS, eigenmaps and spectral clustering. In: *Proc. NIPS*, pp. 177–184 (2003)
12. Mayo-Smith, W.M., Gupta, H., Ridlen, M.S., et al.: Detecting hepatic lesions: the added utility of CT liver window settings. *Radiology* 210, 601–604 (1999)

Evaluation of Medical Image Registration by Using 3D SIFT and Phase-Only Correlation

Zisheng Li¹, Tsuneya Kurihara¹, Kazuki Matsuzaki¹, and Toshiyuki Irie²

¹ Central Research Laboratory, Hitachi, Ltd.,
1-280 Higashi-Koigakubo, Kokubunji-shi, Tokyo 185-8601, Japan

² Department of Radiology, Hitachi General Hospital,
2-1-1 Jonan-cho, Hitachi-shi, Ibaraki 317-0077, Japan
{zisheng.li.fj,tsuneya.kurihara.vn,kazuki.matsuzaki.gv,
toshiyuki.irie.rq}@hitachi.com

Abstract. An effective method for quantitatively evaluating rigid and non-rigid image registration without any manual assessment is proposed. This evaluation method is based on feature point detection in reference images and corresponding point localization in registered floating images. For feature point detection, a 3D SIFT keypoint detector is applied to determine evaluation reference points in liver vessel regions of reference images. For corresponding point localization, a 3D phase-only correlation approach is applied to match reference points and their corresponding points. Distance between the reference points and the correspondences can be used to estimate image registration errors. With the proposed method, users can evaluate different registration algorithms using their own image data automatically.

Keywords: Image registration, evaluation, feature point detection, corresponding point localization.

1 Introduction

Image registration aims to find a spatial transformation that maps points from one image to corresponding points in another image. Medical image registration is fundamentally used in many applications, such as diagnosis, planning treatment, guiding treatment, and monitoring disease progression. Thus, it is necessary to validate whether a rigid/non-rigid registration algorithm satisfies the needs of an image processing application with high accuracy, robustness, and other performance criteria.

The most straightforward method for estimating image registration error is to compare a given registration transformation with a “gold standard” transformation [1], whose accuracy is high. However, the lack of a gold standard prevents any automatic assessment of registration accuracy. Even if individuals trained to interpret medical images are involved in an experimental validation of registration algorithm, it is difficult to provide a method for consistently and accurately assessing individual images. As a result, few attempts have been made to evaluate and compare the performance of image registration algorithms. Even so, two projects that stand out in this regard are

the “Retrospective Image Registration and Evaluation (RIRE) Project” [2] (for evaluating rigid registration accuracy) and the “Non-rigid Image Registration Evaluation Project (NIREP)” [3]. The RIRE project used bone-implanted fiducial markers to obtain a marker-based rigid transformation as the gold-standard transformation. Registration error was measured by calculating the error relative to the gold standard over a set of specified regions. In contrast, instead of providing a gold-standard transformation for accuracy evaluation, NIREP provided four metrics for evaluating the performance of non-rigid image registration. Both projects required manual annotation and segmentation to create evaluation databases, and the evaluation data only included brain images.

In our previous work, for clinical applications such as computer-aided diagnosis and observation of treatment progress [7, 12], rigid and non-rigid image registration algorithms are developed. The present study aims to quantitatively evaluate the accuracy and robustness of these methods by using clinical data, especially abdominal CT images for diagnosis of hepatic tumor. Since we do not have any manually annotated image data as ground truth for this evaluation task, it is necessary to develop an automatic method for evaluating the accuracy of the previously developed registration algorithms. To satisfy that necessity, the following method is proposed. In abdominal CT images, liver vessel regions of reference images are segmented, and feature points in the segmented regions are detected using a 3D SIFT (scale-invariant feature transform) keypoint detector [4]. In floating images after registration, correspondences of the SIFT feature points are searched for and localized using a 3D phase-based image-matching method [5, 13]. By calculating distance between the feature point pairs, it is possible to obtain a numerical-registration error without any manual assessment of the registration algorithm. A brief description of the proposed method is given in Section 2, and the fiducial-point detection and the correspondence matching for evaluating registration accuracy are described in Sections 3 and 4, respectively. Results of experiments on medical image registration are presented in Section 5, and concluding remarks are given in Section 6.

2 Proposed Method

To quantitatively measure registration error, distance between landmarks or regions in reference images and their correspondences in floating images after registration should be accurately estimated. It is therefore necessary to develop an accuracy-evaluation framework that mainly consists of two stages: fiducial point detection and corresponding point matching. The framework of the proposed method is illustrated in Fig. 1, and some data examples generated by different procedures for accuracy evaluation are shown in Fig. 2.

As for landmark detection, to extract reference points, scale-space extrema detection, i.e., keypoint detection in 3D SIFT features [4], is adopted. Since anatomic landmark points of vessel regions are considered as appropriate reference points for evaluating the accuracy of medical image registration [6], liver vessel regions are segmented from our clinical data, and SIFT feature points are extracted from the

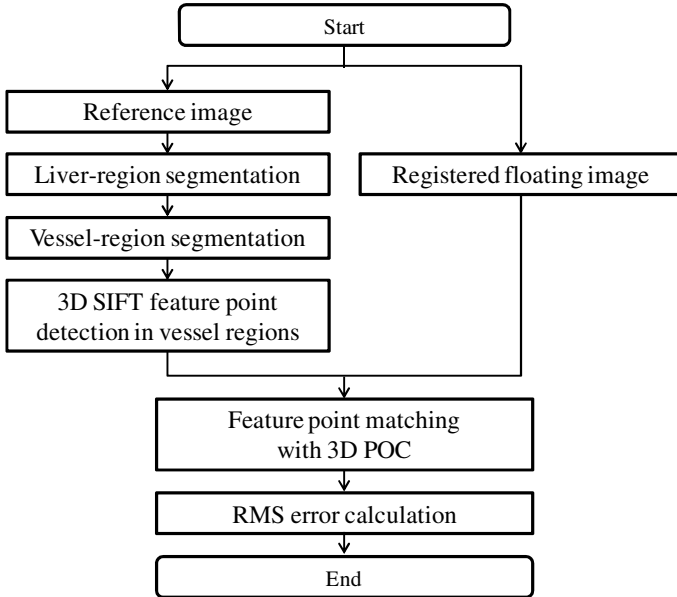


Fig. 1. Framework of evaluation of registration results

segmented regions as reference points. As shown in Fig. 1, the liver region is firstly segmented from the reference image with a previously developed graph-cuts-based method [7]. After that, the vessel region is segmented from the liver region. To detect appropriate feature points for evaluation, the region of interest includes not only liver vessels but also neighboring regions of the vessels. (Such regions are called “vessel regions” for convenience hereafter.) For such regions, high-accuracy segmentation of vessels is not required; instead, a smoothing and thresholding procedure can be used. To extract reference points from the vessel regions automatically, 3D SIFT keypoint detection, which detects local extrema from image pyramids consisting of differences in Gaussian-blurred images at multiple scales, is applied.

After the SIFT feature point detection, instead of SIFT descriptor, a phase-based image-matching algorithm with high accuracy is applied to localize the correspondences. A Fourier transform (FT) of image blocks around the reference points and that of their candidate correspondences is computed, respectively. Phase components of the image-block pairs in the frequency domain are used to estimate locations of the correspondences. The image-block matching method can achieve high accuracy at the sub-voxel level [5]. Even under rotation, the matching method can give good performance by coarse-to-fine and iterative procedures. As a result, it is possible to measure a registration error by using the distance between the SIFT keypoints and their correspondences. Note that although the current implementation

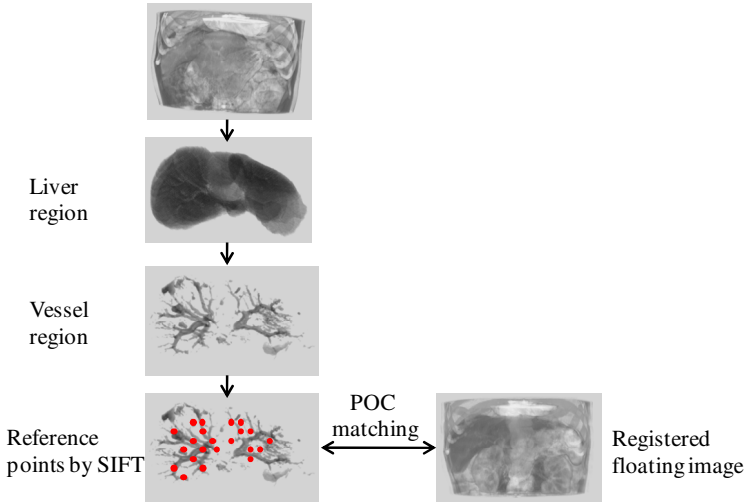


Fig. 2. Examples of fiducial point matching results for registration evaluation

only focuses on feature point matching in liver-vessel regions, the proposed method can be extended to different types and modalities of images, since both the feature-detection and image-matching algorithms can be applied to general image data.

3 Feature Point Detection

3.1 Vessel-Region Segmentation

To detect fiducial points used in registration-accuracy evaluation, first, liver vessel regions are segmented from abdomen images. Thanks to our previous work on liver-region segmentation [7], we assume that segmented liver region was available for this study. A Hessian-based 3D line filter [8] can be used to enhance tubular structures in the liver regions, and major vessels can be segmented by thresholding. However, as mentioned in Section 2, it is necessary to extract fiducial points by using a feature point detector, so the region of interest includes not only liver vessels but also neighboring regions of these vessels. A simple way to obtain such regions is to smooth the liver region and segment regions with threshold T determined by mean μ_I and standard deviation σ_I of the smoothed intensities as

$$T = \mu_I + k\sigma_I, \quad (1)$$

where k is empirically set as $k = 1.0$. Voxels whose intensity exceeds T are considered as the region of interest for feature point detection.

3.2 3D SIFT Feature Point Detection in Vessel Region

To detect fiducial points in reference images, 3D SIFT feature point detector is applied to segmented vessel regions. SIFT feature points can be detected by using difference-of-Gaussian (DoG) images [9]. For a medical image, which is usually a 3D image, DoG images can be computed as follows [4]:

$$D(x, y, z, k^i\sigma) = L(x, y, z, k^{i+1}\sigma) - L(x, y, z, k^i\sigma), \quad (2)$$

where k is a constant multiplicative factor. $L(x, y, z, \sigma)$ is an image $I(x, y, z)$ that is smoothed with a variable-scale Gaussian filter $G(x, y, z, \sigma)$:

$$L(x, y, z, \sigma) = G(x, y, z, \sigma) * I(x, y, z) = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} * I(x, y, z). \quad (3)$$

Such DoG images are used for detecting local extrema at a pyramid level for a certain scale σ . For a voxel v in a DoG image with σ , the intensity of v is compared with those of its 80 neighboring voxels (26 neighboring points at the same σ , and 27 counterparts at the scale of $k^1\sigma$ and $k^{-1}\sigma$, respectively). The voxel with the most or least extreme value of intensity is considered as the feature point. Multi-scale searching parameters are set as $\sigma = \{1.00, 1.26, 1.58, 2.00, 2.52, 3.17, 4.00\}$ and $k = 2.0^{1/3}$. Note that although the DoG images are computed with the whole image, the feature point detection is only performed in the vessel regions.

4 Corresponding Point Localization

To measure distance between the reference points and their correspondences, an image-matching method based on a 3D phase-only correlation (POC) function is applied. The original POC function [10] is calculated with a 2D discrete Fourier transform (2D DFT) to estimate displacement between image blocks, and it is extended to a 3D implementation while maintaining good performance [5].

4.1 Image Matching Using 3D Phase-Only Correlation (3D POC) Function

Given two $N_1 \times N_2 \times N_3$ volumes, $r(n_1, n_2, n_3)$ and $f(n_1, n_2, n_3)$, 3D DFT of the two volumes are $R(k_1, k_2, k_3)$ and $F(k_1, k_2, k_3)$. Normalized cross spectrum $\hat{P}(k_1, k_2, k_3)$ is defined as

$$\hat{P}(k_1, k_2, k_3) = \frac{R(k_1, k_2, k_3) \overline{F(k_1, k_2, k_3)}}{|R(k_1, k_2, k_3) \overline{F(k_1, k_2, k_3)}|} = e^{j\{\theta_R(k_1, k_2, k_3) - \theta_F(k_1, k_2, k_3)\}}, \quad (4)$$

where $e^{j\theta_R(k_1, k_2, k_3)}$ and $e^{j\theta_F(k_1, k_2, k_3)}$ are phase components of $R(k_1, k_2, k_3)$ and $F(k_1, k_2, k_3)$; $\overline{F(k_1, k_2, k_3)}$ denotes the complex conjugate of $F(k_1, k_2, k_3)$. The POC function $\hat{p}(n_1, n_2, n_3)$ between $r(n_1, n_2, n_3)$ and $f(n_1, n_2, n_3)$ is the 3D inverse DFT (3D IDFT) of $\hat{P}(k_1, k_2, k_3)$, and is given by

$$\hat{p}(n_1, n_2, n_3) = \frac{1}{N_1 N_2 N_3} \sum_{k_1, k_2, k_3} \hat{P}(k_1, k_2, k_3) e^{j\frac{2\pi k_1 n_1}{N_1}} e^{j\frac{2\pi k_2 n_2}{N_2}} e^{j\frac{2\pi k_3 n_3}{N_3}}. \quad (5)$$

If two volumes are similar, their POC function gives a distinct sharp peak. If not, the peak drops significantly. The height of the peak can be used as a good similarity measure for image matching, and the location of the peak shows the displacement between the two volumes. When the peak locates between image voxels, the sub-voxel displacement can also be obtained by peak model fitting in good accuracy [10].

In image matching using 3D POC, it is assumed that reference points and their correspondences are at the same voxels initially. Corresponding image blocks around the point-pairs are used to calculate the POC function and estimate displacements between the point-pairs. Moreover, image pyramids of reference and floating images are created, and the block matching is performed in a coarse-to-fine way. The matching procedure starts in the coarsest level while the size of image blocks remain the same in all levels. In our work, the block size is fixed to $32 \times 32 \times 32$ voxels empirically.

When transformation such as rotation is occurred on the image blocks, it is difficult to obtain accurate matching results by a single run of POC matching. In this case, iterative processing is necessary. Firstly, a rigid transformation is obtained by an initial matching with outlier correction. The transformation is applied to the floating image to reduce the rotation between the image pair. Then, block matching is performed again, followed by a resulted transformation. Experimentally, three runs of such iterative processing is able to give good matching results for registration evaluation in our work.

4.2 Combination of 3D SIFT and 3D POC Matching for Registration Evaluation

Since 3D POC can provide accurate image-matching results, we suppose that distance between corresponding landmark pairs can be used to validate accuracy of an image registration. However, one of the main problems in 3D POC matching is the selection of reference points. In [5], the reference points for image matching are determined by CT value empirically, but this determination is not appropriate for registration evaluation. In this work, 3D SIFT keypoint detector is applied to detect fiducial points in vessel regions of reference images. Corresponding points in registered floating images are then searched for and localized using the 3D POC image-matching. We suppose that the distance between the SIFT feature points and their correspondences (localized by 3D POC image matching) is an appropriate measure of image registration error. This supposition was experimentally verified as described in the next section.

5 Experimental Results

To verify the effectiveness of the proposed algorithm for evaluating image registration, the accuracy of the image matching based on a 3D POC function was estimated.

In the rest experiments, the proposed evaluation method was used to evaluate registration results obtained by different registration methods, which include rigid and non-rigid registration algorithms developed in our previous work, and by the rigid and non-rigid registration modules of the Elastix toolkit [11].

5.1 3D POC Matching Error

The accuracy of image-matching algorithm based on a 3D POC function was experimentally validated. In detail, seven sets of simulated data were generated from abdominal CT images of seven patients. Image resolution was 0.548 to 0.665 mm, with 512×512 voxels in the axial plane, and 2.5-mm resolution in the z -direction. In each set of simulated data, reference and floating images were generated from the same CT image. To perform the experiment in an intermodal way, a grey-level transformation was applied to the floating image according to the following equation:

$$f_F = \frac{1 - \cos(2\pi f_R)}{2}, \quad (6)$$

where the range of intensities in f_R is $[0, 1]$. The floating image was then transformed by arbitrary translations and rotations to generate displacements from the reference image. Since the original image pair were the same image, the displacements were known. Fiducial points from every reference image were detected, and correspondences in the generated floating images were localized. The estimated correspondences were compared to the ground truth, and RMS error of the image matching was calculated as follows:

$$\begin{aligned} \text{Err} &= \sqrt{\frac{\sum_i |\mathbf{p}_i - \mathbf{q}_i|^2}{N}} \\ &= \sqrt{\frac{\sum_i \{(p_{x_i} - q_{x_i})^2 + (p_{y_i} - q_{y_i})^2 + (p_{z_i} - q_{z_i})^2\}}{N}}, \end{aligned} \quad (7)$$

where $\mathbf{p}_i = (p_{x_i}, p_{y_i}, p_{z_i})$ is a reference point, $\mathbf{q}_i = (q_{x_i}, q_{y_i}, q_{z_i})$ is a corresponding point, and N is the number of point pairs.

The rotation applied to the floating images was fixed at -2.5 degrees, and the images were then translated from -10 to 10 mm with an interval of 5 mm. The translation was then fixed at 5 mm, and the images were rotated from -5 to 5 degrees with an interval of 2.5 degrees. The RMS matching errors are listed in Tables 1 and 2. It is clear from the tables that the RMS errors are from 0.077 to 0.140 mm, which illustrates that the proposed evaluation method can provide high accuracy for landmark matching and that it is effective for measuring registration errors quantitatively.

Table 1. Image-matching errors of 3D POC function when rotation is -2.5 degrees

Translation (mm)	-10	-5	0	5	10
RMS error:	0.083	0.079	0.077	0.098	0.102
Mean (\pm STD) (mm)	(0.033)	(0.033)	(0.021)	(0.034)	(0.025)

Table 2. Image-matching errors of 3D POC function when translation is 5 mm

Rotation (degree)	-5	-2.5	0	2.5	5
RMS error:	0.140	0.098	0.080	0.097	0.130
Mean (\pm STD) (mm)	(0.023)	(0.034)	(0.028)	(0.037)	(0.020)

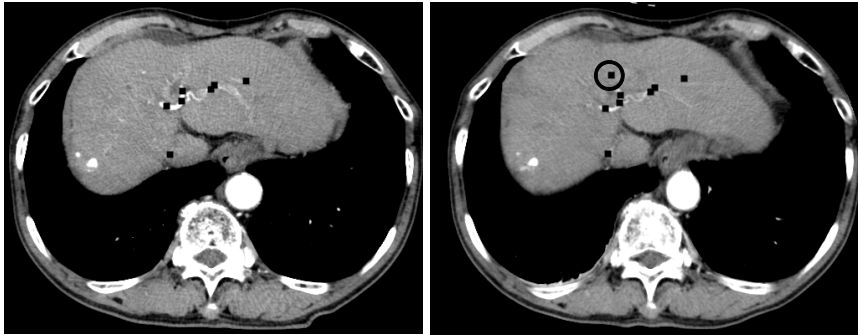
Table 3. Evaluation results for four algorithms

	Rigid	Elastix rigid	Non-rigid	Elastix non-rigid
RMS error:	5.72	8.00	3.29	4.03
Mean (\pm STD) (mm)	(3.23)	(3.60)	(0.92)	(1.61)

5.2 Evaluation of Rigid and Non-rigid Registration Results

The clinical data used in the evaluation experiments are abdominal contrast-enhanced CT images (in arterial phase) from the same patients mentioned in Section 5.1. Since these data are used for observation of disease development and treatment progress in RFA surgery for hepatic tumor, image pairs before and after the RFA surgery, with intervals of one to six months, were selected. In the experiments, four registration methods were evaluated: rigid and non-rigid algorithms developed in our previous work and the similar modules of the Elastix toolkit. Registration parameters of Elastix are set according to the recommendations in the Elastix manual. These four registration methods were applied to the seven image pairs, and registered floating images were created. 3D SIFT feature points in vessel region of each reference image were then detected, and POC-based image matching was performed to estimate locations of correspondences in the registered floating images. The number of SIFT feature points detected from liver-vessel regions was from 200 to 500. The running time of SIFT detection was about 20 to 40 seconds, and that of POC matching was about 20 to 60 seconds. All the evaluation tasks were run on a system with Intel® Core™ i7 3.07-GHz CPU and 12-GB memory without any optimization. Example results of fiducial point matching for evaluating a non-rigid registration are given in Fig. 3 [(a): reference image; (b): floating image after registration]. It is clear that every fiducial point in the reference image has a corresponding point that is accurately located in the registered floating image. Note that the landmark circled in the floating image is the corresponding point in another slice of the reference image.

The RMS error of the four image registration methods applied to a certain image was calculated according to Eq. 7, and the calculated errors are listed in Table 3. Both the two non-rigid algorithms perform a rigid registration as a pre-processing procedure. Here, average errors were only measured in registrations with errors less



(a) Reference image

(b) Floating image after registration

Fig. 3. Examples of fiducial point matching results for registration evaluation

than 12 mm by both the rigid/non-rigid methods and the Elastix rigid/non-rigid modules. As a result, registration results of 5 image pairs given by rigid methods, and those of 6 image pairs given by non-rigid methods were evaluated in our work.

6 Conclusion and Future Work

A method for quantitatively evaluating rigid and non-rigid image registration in an automatic way was developed. With the proposed method, fiducial points for the evaluation are detected by 3D SIFT keypoint detector applied to liver-vessel regions of reference images, and corresponding points in the registered floating images are localized with 3D POC-based image matching. Experimental results show that the proposed method can provide high enough accuracy that the distance between the fiducial/corresponding point pairs can be used to measure image registration error. With the proposed method, users can assess their own image data with different registration algorithms quantitatively and automatically.

Since the proposed method is performed on reference images and registered floating images, in the future, it is necessary to evaluate by an experienced observer that whether the detected matches effectively reflect actual physical correspondences in original image pairs.

References

1. Hajnal, J.V., Hawkes, D.J., Hill, D.L.G.: Medical Image Registration. CRC Press (2001)
2. West, J., et al.: Comparison and Evaluation of Retrospective Intermodality Brain Image Registration Techniques. *Journal of Computer Assisted Tomography* 21(4), 554–566 (1997)

3. Christensen, G.E., Geng, X., Kuhl, J.G., Bruss, J., Grabowski, T.J., Pirwani, I.A., Vannier, M.W., Allen, J.S., Damasio, H.: Introduction to the Non-rigid Image Registration Evaluation Project (NIREP). In: Pluim, J.P.W., Likar, B., Gerritsen, F.A. (eds.) WBIR 2006. LNCS, vol. 4057, pp. 128–135. Springer, Heidelberg (2006)
4. Cheung, W., Hamarneh, G.: N-SIFT: N-dimensional Scale Invariant Feature Transform. *IEEE Trans. Image Processing* 18(9), 2012–2021 (2009)
5. Tajima, Y., et al.: High-accuracy Volume Registration based on 3D Phase-only Correlation. *IEICE Trans. Information and Systems* J94-D(8), 1398–1409 (2011)
6. Charnoz, A., Agnus, V., Soler, L.: Portal Vein Registration for the Follow-Up of Hepatic Tumours. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004, Part I. LNCS, vol. 3216, pp. 878–886. Springer, Heidelberg (2004)
7. Matsuzaki, K., et al.: Development of Computer-aided Diagnosis System using Intelligent Process Manager of Multiple Image Processing Algorithms for Full Automated Detection of Tumors. To be appeared in: CARS 2012 (2012)
8. Sato, Y., et al.: Tissue Classification based on 3D Local Intensity Structures for Volume Rendering. *Transactions on Visualization and Computer Graphics* 6(2), 160–180 (2000)
9. Lowe, D.G.: Distinctive Image Features from Scale-invariant Keypoints. *International Journal of Computer Vision* 60(2), 91–110 (2004)
10. Takita, K., Muquit, M.A., Aoki, T., Higuchi, T.: A Sub-pixel Correspondence Search Technique for Computer Vision Applications. *IEICE Trans. Fundamentals* E87-A(8), 1913–1922 (2004)
11. Klein, S., et al.: Elastix: A Toolbox for Intensity-based Medical Image Registration. *IEEE Trans. Medical Imaging* 29(1), 196–205 (2010)
12. Li, Z., et al.: Efficient Rigid Registration for Medical Images Based on Small Sample Set. *IEICE Technical Report*, MI 111(199), 1–6 (2011)
13. Miyazawa, K., et al.: A Novel Approach for Volume Registration using 3D Phase-Only Correlation. In: Radiological Society of North America (RSNA) 2009, p. 1070 (2009)

Optimal Medial Surface Generation for Anatomical Volume Representations

Sergio Vera^{1,2}, Miguel A. González¹,
Marius George Linguraru³, and Debora Gil²

¹ Alma IT Systems, Barcelona, Spain

² Computer Vision Center, Comp. Science Dep.,
Universitat Autònoma de Barcelona, Spain

³ Sheikh Zayed Institute for Pediatric Surgical Innovation,
Children's National Medical Center, Washington DC, USA
`sergio.vera@cvc.uab.es`

Abstract. Medial representations are a widely used technique in abdominal organ shape representation and parametrization. Those methods require good medial manifolds as a starting point. Any medial surface used to parameterize a volume should be simple enough to allow an easy manipulation and complete enough to allow an accurate reconstruction of the volume. Obtaining good quality medial surfaces is still a problem with current iterative thinning methods. This forces the usage of generic, pre-calculated medial templates that are adapted to the final shape at the cost of a drop in volume reconstruction. This paper describes an operator for generation of medial structures that generates clean and complete manifolds well suited for their further use in medial representations of abdominal organ volumes. While being simpler than thinning surfaces, experiments show its high performance in volume reconstruction and preservation of medial surface main branching topology.

Keywords: Medial surface representation, volume reconstruction.

1 Introduction

One of the most used tools for volumetric shape representation are medial representations. In abdominal imaging, techniques such as M-Reps [9] and CM-Reps [19,18] have shown the potential to model complex anatomical shapes, and are being used in fields such as computational neuroanatomy [20,14], 3D cardiac modelling [15,16], and cancer treatment planning [12].

While other surface representation/parametrization methods model only the external surface of objects [2,3], medial representations can model also the interior of the shape by providing a radial perpendicular coordinate [1] that extends from the medial surface. This allows to parameterize the (possibly diseased) parenchyma of organs, and their internal vascular system, powerful sources of information in organ functionality, analysis and diagnosis. Any medial manifold used to (re)generate anatomical volumes must be simple enough to allow an



Fig. 1. Medial surfaces obtained using a 6-connected neighborhood (a), and a 26 connected neighborhood (b)

easy generation of the radial axis, but complete enough to allow a satisfactory reconstruction of the volume.

Most methods of medial surface computation are based on morphological (ordered) thinning operations on either the original volume or the distance map to its boundary. In any case, they require the definition of a neighborhood set and surface conditions for the removal of simple voxels (those that do not alter the topology if removed) that do not lie on a surface. The complexity of neighborhood definition and surface tests increases exponentially with the dimension of the embedding space [5]. Additionally, small changes in those tests or in the order in which voxels are traversed, generate completely different surfaces (as illustrated in Fig. 1). Surfaces produced with thinning based methods would need to be pruned in order to eliminate spurious branches and manifolds generated due to noise in the volume surface. However, there is no easy way to tell which manifolds can be safely removed without hurting the capability of representation of anatomical structures.

Some authors overcome this limitations using a generic manifold that has to be fitted into the volumetric shape [19]. This limits the number of objects that can be processed to those that can be represented by the topology of the template manifold. Simplified templates do not introduce a high reconstruction error as far as the concavity of the volume boundary keeps low (as for a number of subcortical brain structures [13]). However, anatomies with complex concavity patterns (such as abdominal organs [17]) or pathological shapes with severe deformation can not be captured by a deformed, simplified manifold.

Preserving the medial main branching topology is of prime importance for successfully applying medial representations to any anatomical shape. In order to do so, it would be desirable to generate specific initial manifolds for each anatomical case. This would remove fitting errors in the model and would capture the complete topology of volumes. Recent methods for medial surface generation based on Non Maxima Suppression (NMS) of medialness maps [17] have shown a strong potential to generate surfaces with minimal branching and great reconstruction power. In order to produce complete surfaces, the definition of the medial map is crucial. In [17], authors use a map based on level sets that in spite of giving a normalized response it has two main weak points. On one side, the response is

a step-wise almost binary map that is prone to introduce internal surface holes in the further NMS stage. On the other side, the response significantly drops at surface branches, which, again, might introduce unconnected components.

The present work focuses on the definition of a medial map capable of producing complete medial surfaces reaching a good compromise between simplicity of the medial geometry and its ability to reconstruct the whole anatomical volume. We introduce a medial map based on ridge detectors that combines the advantages of steerable filters and level sets geometry. We call this medial map Geometric Steerable Medial Map, GSM2. A database of liver segmentations generated from an abdominal atlas [11,6] is used as a benchmark to evaluate the accuracy of volumes reconstructed from medial surfaces, as well as the capability for preserving medial main branches. Results show that the proposed GSM2 produces branching topologies related to anatomy concavities that have a reconstruction power higher than thinning approaches.

2 Medial Map Combining Geometric and Steerable Filters

Distance maps are a key element for obtaining medial maps, since, by definition, their maximum values are located at central voxels corresponding to the medial structure. Distance maps can be used directly to generate skeletons (see [10]) but the ridges of the distance map have show superior power to identify medial voxels [17]. In image processing, ridge detectors are based either on level sets geometry or image intensity profiles.

The operator described in [8] defines ridges as lines joining points of maximum curvature of the distance map level sets. It is computed using the maximum eigenvector of the structure tensor of the distance map as follows.

Let \mathbf{V} be the eigenvector of principal eigenvalue of the structure tensor and consider its reorientation along the distance gradient, $\mathbf{V} = (P, Q, R)$, given as:

$$\mathbf{V} = \text{sign}(\langle \mathbf{V} \cdot \nabla D \rangle) \cdot \mathbf{V}$$

for $\langle \cdot \rangle$ the scalar product. The ridgeness measure [8] is given by the divergence:

$$NRM := \text{div}(\mathbf{V}) = \partial_x P + \partial_y Q + \partial_z R \quad (1)$$

The above operator assigns positive values to ridge pixels and negative values to valley ones.

A main advantage is that $NRM \in [-N, N]$ for N the dimension of the volume. In this way, it is possible to set a threshold common to any volume for detecting significant ridges and, thus, points highly likely to belong to the medial surface. However, by its geometric nature, NRM has two main limitations. In order to be properly defined, NRM requires that the vector \mathbf{V} uniquely defines the tangent space to image level sets. Therefore, the operator achieves strong responses in the case of one-fold medial manifolds, but significantly drops anywhere two or more medial surfaces intersect each other. Additionally, NRM responses are not

continuous maps but step-wise almost binary images (see Fig. 2, left). Such discrete nature of the map hinders the performance of the NMS binarization step that removes some internal voxels of the medial structure and, thus, introduces holes in the final medial surface.

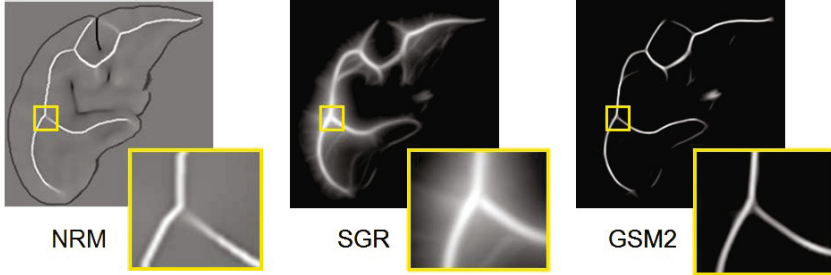


Fig. 2. Performance of different ridge operators. Normalized Ridge Map (left), Steerable Gaussian Ridge (center) and Geometric Steerable Medial Map (right).

On the other side, ridge maps based on image intensity are computed by convolution with a bank of steerable filters. Each filter is defined by 2nd derivatives of (oriented) anisotropic 3D Gaussian kernels:

$$g_{\sigma}^{\Theta} = g_{(\sigma_x, \sigma_y, \sigma_z)}^{(\theta, \phi)} = \frac{1}{(2\pi)^{3/2} \sigma_x \sigma_y \sigma_z} e^{-\left(\frac{\tilde{x}^2}{2\sigma_x^2} + \frac{\tilde{y}^2}{2\sigma_y^2} + \frac{\tilde{z}^2}{2\sigma_z^2}\right)}$$

for $(\tilde{x}, \tilde{y}, \tilde{z})$ the coordinates given by a rotations of angles θ and ϕ that transform the z-axis into the unitary vector $(\cos \phi \cos \theta, \cos \phi \sin \theta, \sin \phi)$. In order to detect sheet-like ridges, the scales are set to $\sigma_z < \sigma_x = \sigma_y$.

The second partial derivative along the z axis constitutes the principal kernel for computing ridges:

$$\partial_z^2 g_{\sigma}^{\Theta} = (\tilde{z}^2 / \sigma_z^4 - 1 / \sigma_z^2) g_{\sigma}^{\Theta}$$

The response of the operator Steerable Gaussian Ridge (SGR) is calculated as the maximum response for a discrete sampling of the angulation:

$$SGR := \max_{i,j} (D * \partial_z^2 g_{\sigma}^{\Theta_{i,j}}) \tag{2}$$

for $\Theta_{i,j}$ given by $\theta_i = \{\frac{i\pi}{N}, i = 1..N\}$ and $\phi_j = \{\frac{j\pi}{M}, j = 1..M\}$.

A main advantage of using steerable filters is that their response provides continuous maps which ensure completeness of the surfaces obtained by NMS binarization. Besides, since they decouple the space of possible orientations for medial surfaces, their response does not decrease at self-intersections (see Fig. 2, left and center). Their main counterpart is that their response is not normalized, so setting the threshold for NMS binarization becomes a delicate issue.

The analysis above shows that geometric and intensity methods have complementary advantages and shortcomings. Therefore we propose combining them into the following Geometric Steerable Medial Map (GSM2):

$$GSM2 := SGR(NMR) = \max_{i,j} (NRM * \partial_z^2 g_\sigma^{\Theta_{i,j}}) \quad (3)$$

The advantages of GSM2 are two-fold. On one hand, steerable filters provide a continuous approximation to NMR semi-discrete maps with a more uniform response at self-intersecting points. On the other hand, because NMR maps have a sharp response at central voxels, GSM2 still provides a highly selective response at ridges. In this manner GSM2 generates medial maps with good combination of specificity in detecting medial voxels while having good characteristics for NMS binarization, which does not introduce internal holes (Fig. 2, right).

3 Validation Experiments

In order to provide a real scenario for the reconstruction tests we have used an atlas of abdominal organs computed by normalized probabilistic models from the registration of 9 subjects [116]. By its higher concavity complexity we have chosen the liver as a source of anatomical volumes. We have applied GSM2 using $\sigma = 0.5$, $\rho = 1$ for computing structure tensors in NMR and $N = M = 8$ orientations for SGR. In order to check the capability of GSM2 for preserving medial main branching topology, we have considered the full surface as well as a simplified surface (labelled $GSM2_S$) obtained by removal of secondary medial branches. For comparison to morphological methods, we have also applied an ordered thinning using a 26-connected neighborhood [10] followed by a pruning (labelled Th_{26P}).

Figure 3 shows the three kinds of medial surfaces considered on a representative liver. In spite of pruning, medial manifolds computed using thinning have a branching geometry more complex than GSM2 and apparently not related to the volume boundary geometry (concave-convex profile). This is not the case for GSM2 surfaces, whose branching topology arises from volume boundary concavities. It follows that its subsequent simplification is better suited for volume convex decomposition, which naturally describe the geometry of objects [7].

Our experiments will focus on evaluating how the pruning affects the generation of reconstructed volumes. Volumes are reconstructed from the computed medial surfaces by applying the inverse medial transform. Comparisons with the original shape are based on volume overlap error (VOE) and maximum symmetric surface distances (MxSD) [4]. Lower metric values indicate better reconstruction capability: VOE provides a global score, while MxSD detects local deviations in the shape of the volume boundary.

Table 1 reports metric scores for each liver. For most livers, thinning is the worst performer in terms of reconstruction power and boundary distortion, while the complete GSM2 is the best method. It is worth noticing that the global volume reconstruction of the simplified GSM2 compares to its complete version

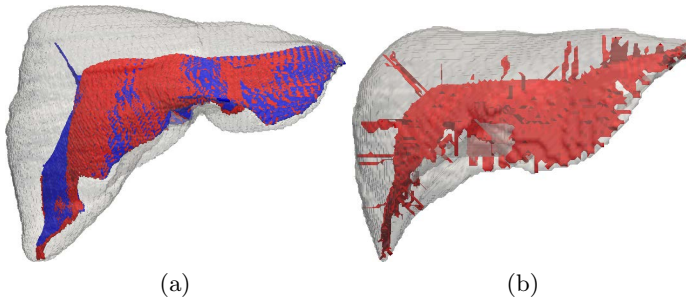


Fig. 3. Medial Manifolds of a healthy liver: full GSM2 (blue surface) and its simplification (red surface), (a) and Th_{26P} , (b)

in all cases. Regarding volume boundaries, we have a noticeable distortion in 3 cases (livers 2, 5 and 7). Distortion appears in the superior liver lobe due to a more prominent concavity in this lobe for the three cases (as illustrated in Fig. 4(a)). This might hinder proper measurements of abnormal or pathological structures. This is not the case for GSM2 complete surfaces as exemplified in Fig. 4. The oversized superior lobe on the right liver is captured by the presence of an unusual medial manifold configuration.

Table 1. Errors in reconstruction VOE and MxSD for each liver

	$GSM2$		$GSM2_S$		Th_{26P}	
	VOE	MxSD	VOE	MxSD	VOE	MxSD
Liver 1	2.51	8.37	2.78	8.37	3.14	12.04
Liver 2	2.22	9	2.85	12.41	2.60	10.68
Liver 3	2.58	8.78	2.92	8.78	3.13	12.04
Liver 4	2.51	9.80	2.55	9.85	2.84	10.77
Liver 5	2.42	4.24	2.83	8.60	2.73	9.64
Liver 6	2.49	9.27	2.71	9.27	2.69	7.49
Liver 7	2.25	4.90	2.69	10.20	2.73	10.48
Liver 8	2.12	10.44	2.36	10.44	2.41	10.77
Liver 9	2.72	9.69	2.98	9.69	3.05	12.04
Mean	2.42	8.28	2.74	9.73	2.81	10.66
Std Dev.	0.19	2.19	0.19	1.23	0.25	1.45

4 Conclusions

In order to provide manageable representations of complex organs, medial manifolds should reach a compromise between simplicity in geometry and capability

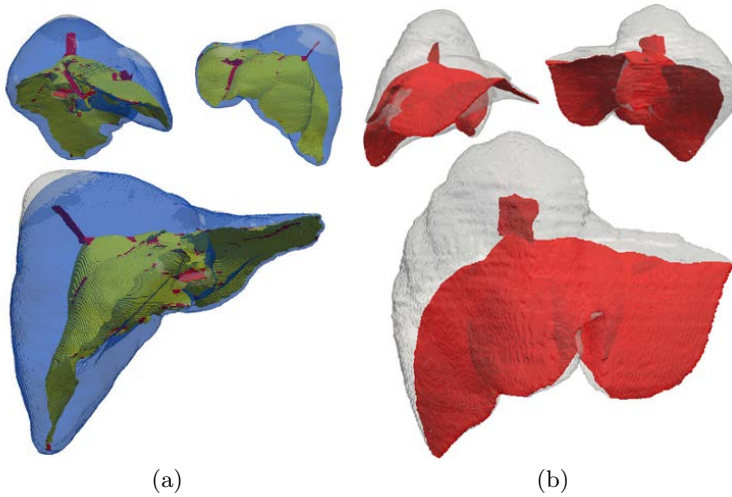


Fig. 4. Impact of medial branching topology: (a) reconstruction error for a healthy liver when using complete GSM2 (white volume and red surface) and simplified GSM2 (blue volume and yellow surface), and detection of unusual lobe for a pathological case, (b)

for restoring the anatomy of the organ. The method presented in this paper allows the computation of medial manifolds resulting in surfaces of greater simplicity than the generated by thinning methods. Although having this minimalistic property, the resulting manifolds can be used to recalculate the original volume with slightly better reconstructions than existing methods.

Our experiments show that our method is preferable to using a generic template since it allows tackling a wider set of shapes which could not be precisely represented by a generic manifold. Any simplification of a medial surface results in a drop in reconstruction quality. This drop in accuracy is hard to relate to the simplification process because the branching topology of thinning-based medial manifolds is not always related to the anatomy curvature (concavity-convexity pattern). A main advantage of GSM2 medial surfaces is that their branches are linked to the shape concavities due to the geometrical and normalized nature of the operator. In this context, GSM2 manifolds can be simplified (pruned) ensuring that the loss of reconstruction power will be minimum.

Future work includes the usage of GSM2 surfaces in the context of shape parametrization, providing a new set of coordinates to each point in the volume.

Acknowledgements. This research has been funded by the Spanish projects TIN2009-13618, CSD2007-00018 and 2009-TEM-00007. The last author has been supported by the Ramon y Cajal Program.

References

1. Blum, H.: A transformation for extracting descriptors of shape. MIT Press (1967)
2. Brechbuehler, C., Gerig, G., Kuebler, O.: Surface parametrization and shape description, pp. 80–89 (1992)
3. Gu, X., Wang, Y., Chan, T.F., Thompson, P.M., Yau, S.: Genus zero surface conformal mapping and its application to brain surface mapping. *IEEE TMI* 23(8), 949–958 (2003)
4. Heimann, T., van Ginneken, B., Styner, M.A., Arzhaeva, Y., Aurich, V.: Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Trans. Med. Imag.* 28(8), 1251–1265 (2009)
5. Lee, T.C., Kashyap, R.L., Chu, C.N.: Building skeleton models via 3-D medial surface axis thinning algorithms. *Graph. Mod. Imag. Process.* 56(6), 462–478 (1994)
6. Linguraru, M.G., Sandberg, J.A., Li, Z., Shah, F., Summers, R.M.: Automated segmentation and quantification of liver and spleen from ct images using normalized probabilistic atlases and enhancement. *Medical Physics* 37(2), 771–783 (2010)
7. Liu, H., Liu, W., Latecki, L.J.: Convex shape decomposition. In: *CVPR*, pp. 97–104 (2010)
8. Lopez, A.M., Lumbreras, F., Serrat, J., Villanueva, J.J.: Evaluation of methods for ridge and valley detection. *IEEE Trans. Pat. Ana. Mach. Intel.* 21(4), 327–335 (1999)
9. Pizer, S.M., Thomas Fletcher, P.: Deformable M-Reps for 3D medical image segmentation. *Int. J. Comp. Vis.* 55(2), 85–106 (2003)
10. Pudney, C.: Distance-ordered homotopic thinning: A skeletonization algorithm for 3D digital images. *Comp. Vis. Imag. Underst.* 72(2), 404–413 (1998)
11. Reyes, M., González Ballester, M.A., Li, Z., Kozic, N., Chin, S., Summers, R.M., Linguraru, M.G.: Anatomical variability of organs via principal factor analysis from the construction of an abdominal probabilistic atlas. In: *IEEE ISBI*, pp. 682–685 (2009)
12. Stough, J.V., Broadhurst, R.E., Pizer, S.M., Chaney, E.L.: Regional Appearance in Deformable Model Segmentation. In: Karssemeijer, N., Lelieveldt, B. (eds.) *IPMI 2007. LNCS*, vol. 4584, pp. 532–543. Springer, Heidelberg (2007)
13. Styner, M., Gerig, G., Lieberman, J., Jones, D., Weinberger, D.: Statistical shape analysis of neuroanatomical structures based on medial models. *Medical Image Analysis* 7(3), 207–220 (2003); *Functional Imaging and Modeling of the Heart*
14. Styner, M., Lieberman, J.A., Pantazis, D., Gerig, G.: Boundary and medial shape analysis of the hippocampus in schizophrenia. *Medical Image Analysis* 8(3), 197–203 (2004)
15. 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. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part II. LNCS*, vol. 5242, pp. 766–773. Springer, Heidelberg (2008)
16. Sun, H., Frangi, A.F., Wang, H., Sukno, F.M., Tobon-Gomez, C., Yushkevich, P.A.: Automatic Cardiac MRI Segmentation Using a Biventricular Deformable Medial Model. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) *MICCAI 2010, Part I. LNCS*, vol. 6361, pp. 468–475. Springer, Heidelberg (2010)
17. Vera, S., Gil, D., Borràs, A., Sánchez, X., Pérez, F., Linguraru, M.G., González Ballester, M.A.: Computation and Evaluation of Medial Surfaces for Shape Representation of Abdominal Organs. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) *Abdominal Imaging 2011. LNCS*, vol. 7029, pp. 223–230. Springer, Heidelberg (2012)

18. Yushkevich, P.A.: Continuous medial representation of brain structures using the biharmonic PDE. *NeuroImage* 45(1), 99–110 (2009)
19. Yushkevich, P., Zhang, H., Gee, J.: Continuous medial representation for anatomical structures. *IEEE Trans. Medical Imaging* 25(12), 1547–1564 (2006)
20. 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)

Registration of Free-Breathing Abdominal 3D Contrast-Enhanced CT

Blandine Romain^{1,2,5}, Véronique Letort¹, Olivier Lucidarme⁴,
Florence d'Alché-Buc^{3,5}, and Laurence Rouet²

¹ MAS Lab, Ecole Centrale Paris, Chatenay-Malabry, France

{`blandine.romain,veronique.letort`}@ecp.fr

² Medisys, Philips Research, Suresnes, France

`laurence.rouet@philips.com`

³ INRIA-Saclay, LRI CNRS 8623, Orsay, France

⁴ AP-HP, Hospital La Pitié-Salpêtrière, Paris, France

`olivier.lucidarme@psl.aphp.fr`

⁵ IBISC, University of Evry, Evry, France

`florence.dalche@ibisc.fr`

Abstract. CT perfusion imaging is used for the follow-up of abdominal tumors. A specificity of our work is that patients are breathing freely during image acquisition (5 minutes). We propose an automatic 3D image registration to compensate for respiratory motion. The registration is computed in two main steps: global translation in the z-direction and 3D multiresolution blockmatching. Within this algorithm, the choice of similarity measure largely determines the algorithm robustness in presence of intensity shifts due to contrast diffusion. We exploit a modified entropy-based similarity measure to improve the quality of registration. We also propose two relevant criteria allowing to quantify the registration quality: one based on the gradients of difference images and one based on the smoothness of enhanced-intensity curves.

Keywords: Free-breathing, spatio-temporal registration, DCE-CT, evaluation criteria.

1 Introduction

Functional imaging has gained attention for oncology therapy. It is based on the acquisition of time sequences combined with contrast injection. Long acquisitions (up to 5 minutes) allow to study the complete dynamics of diffusion, including the effects of tissue permeability. However, it is possible to get relevant information only if data are correctly registered. Two main approaches are possible to deal with respiratory motion, sequential breath hold [1] or free-breathing. In the present study, free breathing is selected, since a less stressed patient usually provides sequences with more regular motion. The current study focuses on tumors in liver or kidney using dynamic enhanced CT.

Methods for registration of images acquired with free-breathing are being developed essentially for contrast-enhanced MRI. Wollny et al. [7] exploit the quasi

periodicity of respiratory cycle for myocardial perfusion. Li et al. [2] explicitly estimate the respiratory cycle within abdominal perfusion. In both cases, an important assumption is a high time resolution.

In CT, it is also possible to have high time resolution, but this is combined with a short time of acquisition (2 minutes) [4], and this does not allow to extract permeability. For CT perfusion, with the objective of extracting adequate data for functional analysis, a longer time of acquisition is required which implies, mostly for dose considerations, a low time resolution. The second constraint is a limited imaging field of view in z-direction. To the best of our knowledge, no previous studies have investigated image registration on free-breathing with abdominal dynamic contrast-enhanced CT (DCE-CT) for long acquisitions. In that context, the aim of this paper is, in a first part, to propose the complete pipeline of the intensity-based methods that we have proposed to compensate free respiratory motion. The second main contribution of this study is the definition and evaluation of a dedicated similarity measure adapted from the difference entropy (DE), which is compared to the classically used DE [3]. Finally, we propose the definitions of a temporal and a spatial criteria, in order to objectively evaluate the quality of the registration and to compare the results obtained when testing the different similarity measures.

2 Materials and Methods

2.1 Data Acquisition

39 data acquisitions were performed with a 256 slices CT (Brilliance iCT 256, Philips Healthcare, The Netherlands) on 14 patients. All patients gave informed consent. The dynamic CT protocol consists in acquiring 48 volumes every 2.5 seconds then 18 volumes every 10 seconds (80 kV, 80 mAS, rotation time of 0.33 seconds, dose of iodine 80 cc). With 8 cm detector coverage in z-direction, the CT scanner allows for 5.5 cm effective z-coverage in a single rotation with a 3D axial cone beam correction. This coverage is large enough to keep most of the lesion in image volume. In the axial plane, data are reconstructed with a pixel size of 0.68 mm \times 0.68 mm [5]. Reconstructed volumes are expressed in Hounsfield Units (HU). Image intensities can thus range from -1024 (air) to 2000 HU (bone).

2.2 Spatio-Temporal Registration

The main challenges for registration in these conditions are:

- Limited volume height: data acquisition covers about 6 cm in z-direction, while the amplitude of respiratory motion may reach 2 to 3 cm, in the same z-direction. This means that the whole liver or kidneys are not imaged. Consequently, full organ tracking is not possible and local approaches have to be used.

- Contrast agent injection: intensities of a given tissue vary with time (figure [1](#)). The choice of a similarity measure which is robust to contrast-induced intensity variations is essential in our method.
- Free breathing and low time resolution: since patients breath freely, a first consequence is that acquisitions correspond to different instants in respiratory cycle and a prediction using a respiratory model is not possible. A second consequence is the existence of strong intensity variations between two consecutive volumes due to contrast arrival and low time resolution (2.5 or 10 seconds). In such a context, the problem of intensity differences is present even between successive pairs of volumes (I_t onto I_{t-1}) and registering these pairs of volumes would lead to unnecessary errors accumulation. So, it is preferable to align all volumes onto a common reference volume I_r . Since the majority of image volumes are enhanced, an intuitive solution would be to choose a reference volume I_r with contrast ($t \approx 30$). However, we observed that the tumor is most often well centered in the first volume ($t = 0$), which is why we selected it as the reference volume I_r .

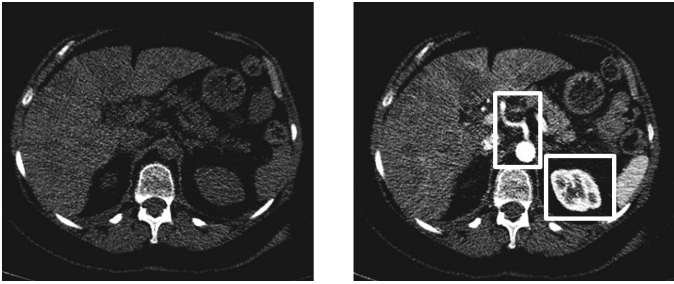


Fig. 1. Image data at time 0 (left) and 15 seconds after injection (right). On the right figure: kidney is enhanced (white square) and aorta with hepatic artery are enhanced (white rectangle, middle).

In order to deal with these challenges, we present an adapted spatio-temporal registration pipeline. Each reconstruction provides a dynamic sequence $DS_I = (I_t, t \in (1, \dots, T))$, where I_t is a 3D image acquired at time t and T is the number of acquisition times.

Our registration approach involves two main steps, followed by a step of regularization and warping:

1. **Global z-Translation** : since the main motion is in the z-direction, the first step consists in a global evaluation of the z-translation. To improve this initial step, thresholds are applied to exclude air (~ 1024 HU) and intense contrasts (≥ 600 HU), such as the contrast of the static aorta in z-direction. We consider all pixels which intensity is between -800 HU and 600 HU.

Since we are considering one global translation for the whole masked volume at this point, we verified that the local contrast diffusion was not an important disturbance, and that the sum of square of differences (SSD) was the most appropriate metric (compared to entropy and mutual information, results not shown).

2. **Multi-Resolution Blockmatching** [6]: initialized by the z-translation found in the first step, this 3D registration method is computed with block sizes of $(11 \times 11 \times 7 \text{ mm}^3)$ at the highest resolution. It provides, for each block, motion vectors corresponding to translations in x, y and z directions.
3. **Regularization**: a regularization step is necessary to smooth the motion vector fields. A gaussian filter is computed in 3D ($\sigma = 1.2$).
4. **Warping**: warping consists in reconstructing new image sequences (J_1, \dots, J_T) with respiratory motion compensation. A trilinear interpolation is used.

2.3 Masking

In order to improve the algorithm robustness and its computing efficiency, masked volumes are computed. For each volume, a mask is defined by excluding spine, which have a different motion from the rest of the abdomen, and background, which does not give useful information.

Our mask consists of two 3D zones with z-translation invariance. These two zones can therefore be represented in an axial plane of arbitrary z-value.

On each axial plane, a threshold of 500 HU is first applied to extract the bones: the largest connected component is the spine. Similarly, the largest connected component of pixels of intensity values lower than -500 HU are considered to belong to the background (fig. 2).

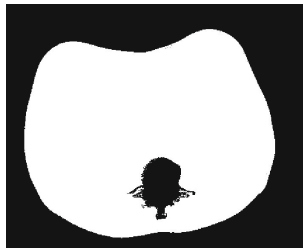


Fig. 2. Mask example: white zone is the valid region, excluding spine and background

2.4 Similarity Measures for Multi-resolution Blockmatching

To locally register the volumes I_r and I_t , our approach is based on a similarity measure which is potentially robust to intensity variations, namely DE.

The entropy H of a discrete random variable X with probability mass function $p(X)$ is defined as:

$$H(X) = \mathbb{E}(-\log p(X)). \quad (1)$$

The difference between I_r and I_t , noted I_{r-t} , is used to evaluate DE. The classical definition is based on the normalized histogram $hist(I_{r-t})$ of I_{r-t} with B bins:

$$DE = - \sum_{j=1}^B hist(I_{r-t})_j * \log(hist(I_{r-t})_j). \quad (2)$$

DE is classically used for registration, in particular in contrast enhanced echocardiography [3].

However, it is known that the quality of registration is very dependent on the number of bins B . So, our approach is to approximate DE to have a more dedicated and robust similarity measure. We can define H directly in the image domain Ω as:

$$H = \frac{1}{card(\Omega)} \sum_{i=1}^{card(\Omega)} -\log p_i. \quad (3)$$

with p_i the intensity of pixel i and $card(\Omega) = N_x * N_y * N_z$ the number of pixels in Ω , N_x in x dimension, N_y in y dimension and N_z in z dimension.

In addition, assuming that the pixel intensities in each block are distributed following an univariate Gaussian distribution with mean μ and variance σ , DE_G is expressed as:

$$DE_G = \log(\sigma) + \frac{1}{card(\Omega)} \sum_{i=1}^{card(\Omega)} \frac{(p_i - \mu)^2}{2\sigma^2}. \quad (4)$$

We can then compare our proposed similarity measure DE_G and the classical one DE . To this end, we define two evaluation criteria.

2.5 Evaluation Criteria

In order to evaluate the quality of registration for each similarity measure, we define two criteria, related to the temporal and spatial dimensions of the method.

Temporal Evaluation Criterion: We propose a new evaluation criterion based on curve smoothness. The quality of the registration is assessed by the smoothness of the time curve of mean intensities after registration (figure 3). Indeed, after registration, time-intensity curves should reflect the contrast intake only, and not the respiratory motion. Since the presence of kidney in a given axial slice is highly sensitive to respiratory motion, a region of interest (ROI) is defined within the kidney. Mean intensities \bar{I}_t of any image (I_t) before registration and \bar{J}_t of the same image after registration are calculated within ROI, for each acquisition time t and plotted, see fig. 3 for an illustrative example.

The two curves are independently filtered by a median time-filter with a window span of 3 points, noted $M(I_t, 3)$ and $M(J_t, 3)$. Therefore, the sum of the absolute difference between the smoothed curve and the original curve can be

used as indicator of the smoothness of the enhancement curve. Thus, we define SC as:

$$SC = \frac{\sum_{t=1}^T |\bar{J}_t - M(J_t, 3)| - \sum_{t=1}^T |\bar{I}_t - M(I_t, 3)|}{\sum_{t=1}^T |\bar{I}_t - M(I_t, 3)|}. \tag{5}$$

The higher SC is, the better quality of registration is.

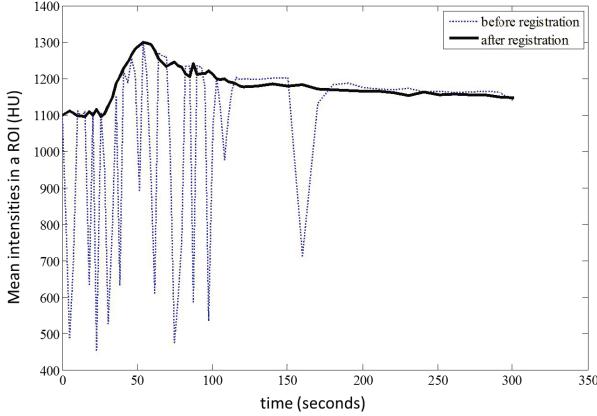


Fig. 3. Mean intensities inside ROI on an axial slice during time of acquisition before registration \bar{I}_t (dash line) and after registration \bar{J}_t (solid line)

Spatial Evaluation Criterion: the usual metric of SSD was not used in our study since a change in contrast intensity would fool the evaluation. In order to remove variations due to contrast arrival, we decided to use the gradient of the difference images. We computed two gradient image volumes $G_{I_t-I_r}$ and $G_{J_t-I_r}$, and their associated norm $\|G_{I_t-I_r}\|$ and $\|G_{J_t-I_r}\|$. Finally, we compute the spatial evaluation criterion, noted GC , as in equation 6. This definition is based on the idea that, after registration, the gradients of difference images should be lower than before registration. So, a high GC shows a better quality of registration.

$$GC = \frac{1}{T} \sum_{t=1}^T \frac{\|G_{I_t-I_r}\| - \|G_{J_t-I_r}\|}{\|G_{I_t-I_r}\|}. \tag{6}$$

G_I is defined as the classical gradient of an image I :

$$G_{I(x,y,z)} = \sqrt{\nabla_X(I(x,y,z))^2 + \nabla_Y(I(x,y,z))^2}. \tag{7}$$

where $\nabla_X(I)$ is the gradient of I in x direction and $\nabla_Y(I)$ is the gradient of I in y direction.

From this gradient image, the L_2 norm is derived:

$$\|G_I\| = \sum_{z=1}^{N_z} \sqrt{\sum_{(x,y)=(1,1)}^{(N_x,N_y)} G_I(x,y,z)^2}. \quad (8)$$

3 Results

Fig. 4 shows the results of registration using a checkerboard. After registration, frontiers of organs are better aligned.

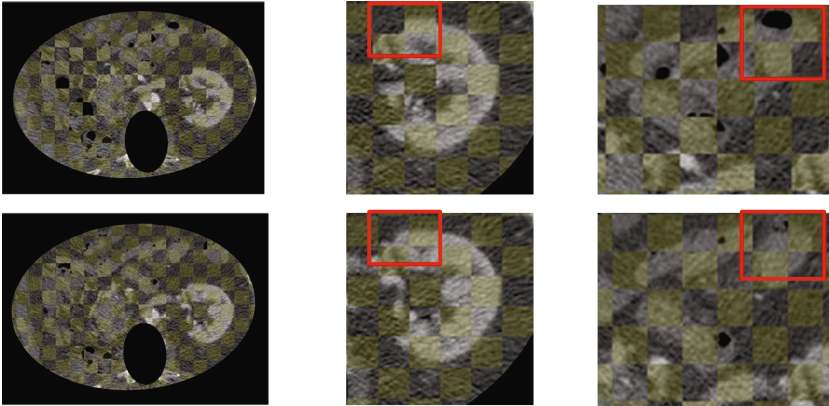


Fig. 4. On the top: before registration; on the bottom: after registration. At left: Checkerboards between image of reference (yellow) and current image (gray). At middle: zoom on kidney, which is better aligned after registration than before, especially in the red squares. At right, zoom on jejunum.

After global z-translation, we apply our algorithm of multiresolution block-matching with DE and our modified similarity measure DE_G . In figure 5, results show that our proposed measure DE_G is better adapted as similarity measure, according to SC and GC criteria. Note that we tested the assumption of gaussian distribution of pixel intensities in each block. It obviously depends on the size of blocks and their location (blocks that cover two different regions of the image have less normal distribution than blocks that include only homogeneous regions). Our algorithm was nevertheless robust even for blocks that did not verified this assumption.

4 Conclusion

We have presented a full setup to provide automatic registration of 4D DCE-CT sequences, using an algorithm that consists of global z-translation and multiresolution blockmatching. One main challenge, related to contrast arrival, has been

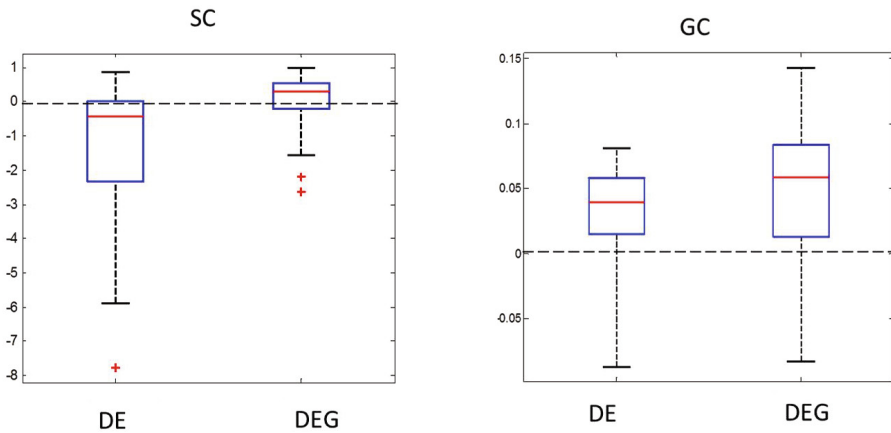


Fig. 5. Values of the two criteria, namely intensity curve smoothness SC (left) and gradient criterion GC (right) for the two similarity measures DE and DE_G . These plots show that the proposed similarity measure DE_G improves the evaluation criteria on the all 39 data acquisitions.

specifically evaluated in the step of multi-resolution block-matching. In this part, we presented an adapted similarity measure (DE_G), which was found to be more adapted to registration on free-breathing abdominal 4D contrast enhanced CT than DE .

Our mid-term objective is to provide a prototype that could be used by clinicians. Therefore, the computation time was one of our main constraints. Our method is relatively efficient. To register one image volume ($512 \times 512 \times 22$), it takes about 3 seconds (Intel(R) Core(TM) i5-2450M CPU, 2,5 GHz). Such a performance is encouraging for a future clinical use.

A possible further refinement of our work could be to consider changing of similarity measures with time, according to the diffusion phases of the contrast.

Acknowledgements. We thank Shrif Makram-Ebeid and Vincent Auvray for their discussions and Philippe Coulon for his implication.

References

1. Koh, T., et al.: Dynamic contrast-enhanced ct imaging of hepatocellular carcinoma in cirrhosis: feasibility of a prolonged dual-phase imaging protocol with tracer kinetics modeling. *Eur. Radiol.* 19(5), 1184–1196 (2009)
2. Li, Z., Caan, M.W.A., Ziech, M.L., Stoker, J., van Vliet, L.J., Vos, F.M.: 3D Non-rigid Motion Correction of Free-Breathing Abdominal DCE-MRI Data. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) *Abdominal Imaging 2011*. LNCS, vol. 7029, pp. 44–50. Springer, Heidelberg (2012)

3. Rodrigues, E.P., et al.: Evaluation of similarity measures in contrast enhanced echocardiography motion detection and registration. In: Proc. Computers in Cardiology, pp. 773–776 (2009)
4. Romain, B., Lucidarme, O., Dauguet, J., Mul, S., Souedet, N., Chenoune, Y., Guibal, A., Delzescaux, T., Frouin, F.: Registration and functional analysis of ct dynamic image sequences for the follow-up of patients with hepatic tumors undergoing antiangiogenic therapy. *IRBM* 31(5-6), 263–270 (2010)
5. Romain, B., et al.: Optimisation of ct reconstruction for the registration of ct liver perfusion sequences. In: Proceedings of SPIE, vol. 8314(1) (February 2012)
6. Tzovaras, D., et al.: Evaluation of multiresolution block matching techniques for motion and disparity estimation. *Signal Processing: Image Communication* 6(1), 59–67 (1994)
7. Wollny, G., Ledesma-Carbayo, M.J., Kellman, P., Santos, A.: Exploiting quasiperiodicity in motion correction of free-breathing myocardial perfusion mri 29(8), 1516–1527 (2010)

Non-newtonian Blood Flow Analysis for the Portal Vein Based on a CT Image

Harvey Ho¹, Adam Bartlett², and Peter Hunter¹

¹ Bioengineering Institute, University of Auckland, 70 Symonds Street, City Campus,
Auckland 1010, New Zealand

`{harvey.ho,p.hunter}@auckland.ac.nz`

² Department of Surgery, University of Auckland, School of Medicine, 85 Park Road,
Grafton, New Zealand

`a.bartlett@auckland.ac.nz`

Abstract. In this paper we perform a Newtonian and a non-Newtonian blood flow analysis for a patient-specific portal vein (PV), which was digitized from a CT image. The non-linear relationship between the shear stress and shear rate was simulated using a Carreau model. We found that, under normal physiological conditions, the computed data from the non-Newtonian model was only marginally different from that of the Newtonian model. However, when the portal flow was severely reduced (e.g., 10% of its normal value), the difference between the two models was significant. Hence we suggest that the Newtonian model is a good approximation for portal flow in physiological conditions whereas a non-Newtonian model should be used in pathological conditions when the very low flow rate induces a much higher blood viscosity.

Keywords: Blood flow, computational modeling, portal vein, computed tomography.

1 Introduction

In healthy adult human beings, about 30% ($\sim 1.5\text{L}/\text{min}$) of total blood flow volume perfuses the liver through a hepatic artery and a portal vein (PV). In particular, the PV is the only pathway to deliver $2/3$ ($\sim 1\text{L}/\text{min}$) of this large volume of nutrient-borne but poorly-oxygenated blood from the intestine, pancreas and spleen into the liver [3].

Anatomically the PV bifurcates into the left PV and right PV which supply their respective lobes. This can be appreciated from tomography images after the portal veins are enhanced by a contrast agent, e.g., as shown in Fig. 1(a). The CT image alone, however, does not provide any quantitative information about blood flow such as its flow velocity and pressure. In contrast, phase contrast MRA is able to yield 3D flow quantities which may be used for portal flow analysis [3, 1]. However, a MRI scan is costly, and the resolution of the current PC MRA ($\sim 0.8\text{mm}$) is still not high enough to capture complex 3D flow patterns.

Computational modeling provides another means to evaluate blood flow. This has been made possible through advancing computational fluid dynamics (CFD),

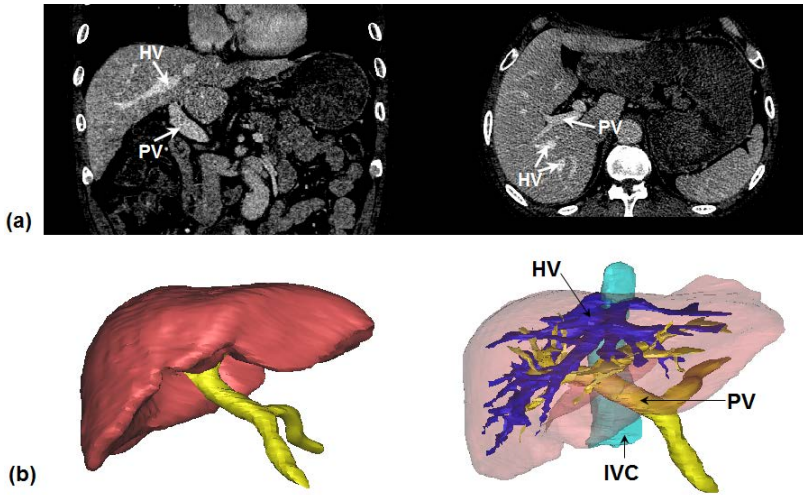


Fig. 1. (a) Portal vein viewed from a CT image; (b) digitized liver surface and vasculature

which, when combined with contemporary image scanning techniques, can be used for flow analysis in a clinical context. Nowadays, general-purpose commercial CFD packages are able to simulate blood flow in patient-specific geometries. While this utility is becoming more frequently used in biomedical research, care should be taken towards the subtle hemodynamic features that might be neglected but can be physiologically significant. For instance, up to this day, flow simulation for the portal vein has been mainly performed assuming the blood as a Newtonian fluid [4], whereby the shear stress (τ) has a linear relationship with shear rate ($\dot{\gamma}$) and its viscosity (μ) is a constant. However, blood is a non-Newtonian fluid and it exhibits shear-thinning properties [7]. A question thus arises as to whether the non-Newtonian blood properties have an impact on blood simulations in the PV.

This purpose of this paper is to address this question by comparing the flow dynamics in both Newtonian and non-Newtonian models, based upon the vascular geometry digitized from a CT-image. We briefly introduce the procedure for vascular model construction, flow analysis, and present the simulation results.

2 Methods

2.1 Medical Imaging

We retrospectively studied the CT image (GE LightSpeed) of a male patient. The spatial resolution of the image was $0.879 \times 0.879 \times 0.625$ mm. The image, shown in Fig. 1(a), enhances the portal and hepatic veins. Using a commercial MIMICS software (Materialise, Leuven, Belgium) we segmented the liver and intrahepatic

PV and HV trees (Fig. 1(b)). Also segmented are the superior mesenteric vein (SMC) and the splenic vein (SV), which merge into the PV.

To facilitate 3D flow simulation, the portal veins downstream the second generation were discarded because the image resolution was not high enough to conduct an accurate 3D vascular surface-reconstruction (see Fig. 2). The surface mesh of the PV, SMV and SV were imported into a computational grid generation software ANSYS ICEM (ANSYS Inc.). The number of generated tetrahedral elements was about 160K using a robust Octree method.

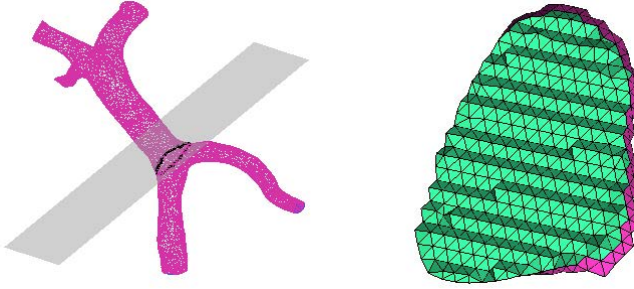


Fig. 2. Computational grid generation: tetrahedral grid is shown at the cross-section sliced by the cut plane

2.2 Mathematical Modeling for Blood Flow

Governing Equations. The governing equations i.e. the mass and momentum conservation equations for blood flow can be written in a compact form:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{v} = 0, \quad (1)$$

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = -\nabla p + \nabla \cdot \boldsymbol{\tau} + \rho \mathbf{g}, \quad (2)$$

where \mathbf{v} , ρ , $\boldsymbol{\tau}$ are the blood flow velocity, blood density and stress tensor, respectively. The last term in Equation (2) $\rho \mathbf{g}$ represents gravity and is often dropped in arterial flow since it is small compared with the pressure force. If the flow is steady, as will be simulated for the portal vein, the first term at the LHS of Equations (1) and (2) can also be dropped.

For Newtonian blood flow, a linear relationship between the stress $\boldsymbol{\tau}$ and strain rates $\dot{\boldsymbol{\gamma}}$ can be expressed as:

$$\boldsymbol{\tau} = \mu \dot{\boldsymbol{\gamma}}, \quad (3)$$

where the blood viscosity μ is a constant ($\mu = 0.00345 Pa \cdot s$).

Non-newtonian Blood Models. The Newtonian blood model used in Equation (3) is only accurate when the flow shear rate ($\dot{\gamma}$) is sufficiently high (e.g., when $\dot{\gamma} > 100s^{-1}$), whereas blood is a shear thinning fluid and its viscosity is much higher in low shear rates (see data from [8]).

An overview of this topic (blood rheology) can be found in [2] and some non-Newtonian models (e.g., the Carreau, Casson and Power Law models) are presented in [7]. In this work we need to determine whether the Newtonian model is a plausible approximation. To that end we implemented the Carreau model into the solver, which represents the relationship between the viscosity and shear rate as:

$$\mu = \mu_{\infty} + (\mu_0 - \mu_{\infty})[1 + (\lambda\dot{\gamma})^2]^{(n-1)/2}. \tag{4}$$

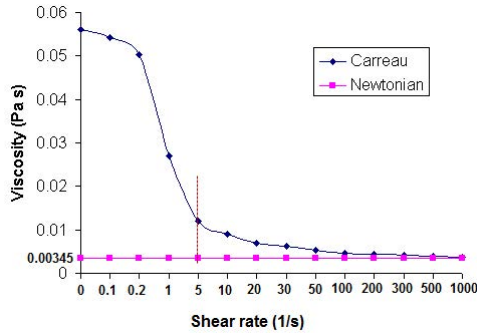


Fig. 3. Viscosity vs shear rate: a comparison of non-Newtonian model (Carreau) and Newtonian model

The parameters μ_0 , μ_{∞} , n and λ are obtained by fitting Equation (4) with experimental data [7]. We adopted their values from [7]: $\lambda = 3.313s$, $n = 0.3568$, $\mu_0 = 0.056Pa \cdot s$ and $\mu_{\infty} = 0.00345Pa \cdot s$, the resulting viscosity-shear rate (μ - $\dot{\gamma}$) curve is plotted in Fig. 3. Note, that the blood viscosity becomes substantially higher when $\dot{\gamma} < 5s^{-1}$, and it tends to be equivalent to the Newtonian viscosity (μ_{∞}) in high shear rates ($> 100s^{-1}$).

Numerical Methods. Equations (1) and (2) were integrated over each of the small elements (also known as control volumes) of Fig. 2 to yield discretized equations. A commercial finite-volume-based flow code CFX 11.0 (ANSYS Inc., Canonsburg, PA) was used to solve these equations numerically. The portal flow was considered as steady, as revealed by ultrasonic measurements [4]. The inflow velocity boundary conditions, 20 cm/s and 30 cm/s, were prescribed at SV and SMV, respectively. These data were adopted from literature [4,9]. In addition, a zero pressure was imposed at the outlet to allow free outflow. The wall of portal

vein was simplified as rigid because of the steady venous flow would not cause large wall displacements.

3 Results

3.1 Newtonian Flow Model

The computation was performed on a desktop computer (Intel Core 2.4GHz). The computational results such as the flow velocity and pressure were post-processed in Fig. 4. We observe that the average flow velocity in PV is about 0.15m/s, which translate into flow rate is about 20 mL/s or 1.2L/min. This is consistent with the flow rate mentioned in Introduction. Also observable are the helical flow pattern (indicated by the arrow) in PV developed after the merging point of the SMV and SV.

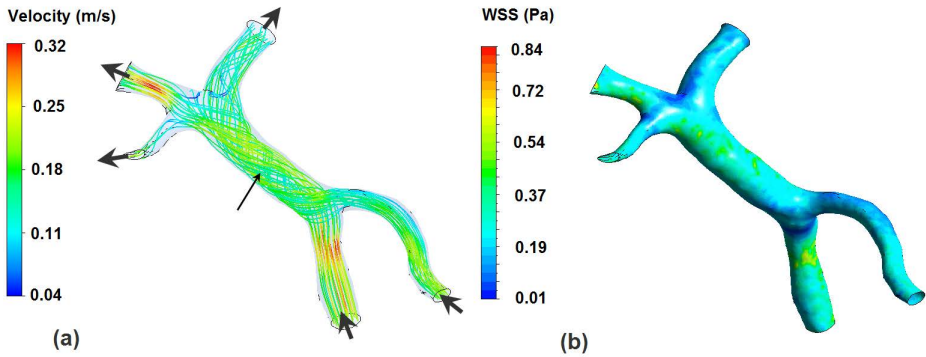


Fig. 4. Post-processing of CFD results: (a) Velocity streamline: helical flow was developed after the merging point of SMV and SV; (2) WSS distribution on the wall

Fig. 4(b) shows the wall shear stress (WSS) distribution on the wall, which ranges from 0.2 to 0.4 Pa, or 0.0015-0.003mmHg. This agrees with the simulation reported in [4]. A visualization was also made for the shear rate, which is shown in Fig. 5. It is evident that the shear rates at very limited regions actually are lower than 5s^{-1} where the blood viscosity is substantially higher (refer to Fig. 3). This might suggest the soundness of employing Newtonian model for flow analysis. Nevertheless, we performed a non-Newtonian flow analysis as below.

3.2 Non-newtonian Models for Normal PV Flow

With the same boundary conditions, but a different transport property (viscosity) as per Equation (4), a non-Newtonian flow analysis was performed and the results are shown in Figs. 6 and 7.

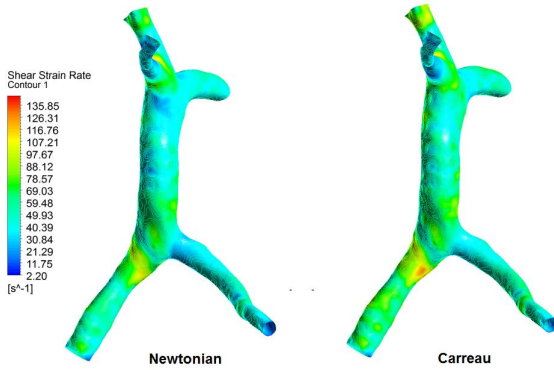


Fig. 5. Shear rate distribution on the portal vein surface: a comparison of the Newtonian and non-Newtonian model

Fig. 6 shows the flow vector profile on a cross-section of the PV. No substantial difference between the Newtonian and Carreau models were noticed, although the velocity vector is slightly smaller for the latter. Fig. 7(a) shows the velocity vector from a different view, which reveals the helical flow pattern. It can be seen that there are only slight difference between the Newtonian and non-Newtonian models. Fig. 7(b) shows the pressure gradient on the cross-section. Again the differences between the Newtonian and non-Newtonian models are only marginal.

3.3 Non-newtonian Models for Low PV Flows

The above simulations were made based on the assumption that the PV flow was normal, i.e., $\sim 1\text{L}/\text{min}$. In pathological conditions, the PV flow can be very low due to a high hepatic resistance, which usually causes portal hypertension. In this scenario the high flow shear rate shown in Fig. 5 is no longer realistic, but according to the plot in Fig. 3 will render significant computational differences between the Newtonian and non-Newtonian models.

Two virtual flow experiments were performed to verify this hypothesis, whereby the PV flow was reduced to $1/3$ and $1/10$ of its normal value. The resulted WSS visualization is shown in Fig. 8. It can be seen that significant differences indeed exist between Newtonian and non-Newtonian models, in particular for the $1/10$ case: the WSS predicated by the Carreau model was substantially higher ($\sim 100\%$) than the Newtonian model.

4 Discussion

A good understanding of blood flow dynamics in the portal vein is important for hepatic physiology and pathology as it is the only vascular pathway to deliver

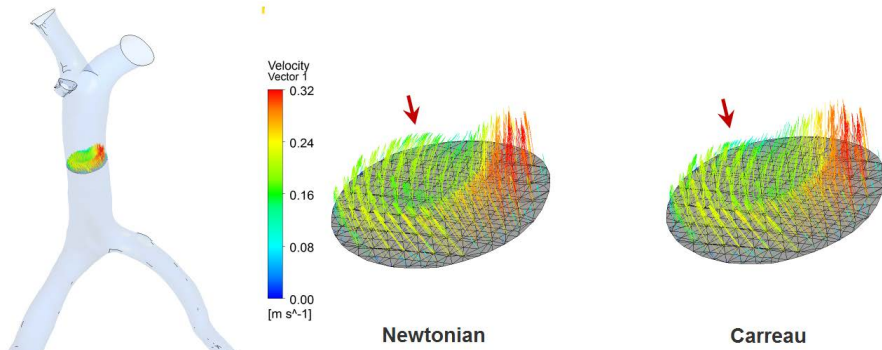


Fig. 6. 3D vector visualization on a cross-section of the PV. No substantial difference between the Newtonian and Carreau models were noted, although the velocity vector is slightly smaller for the later.

nutrients to the liver. Diseases of the portal vein, for instance, portal hypertension which is caused by increased resistance in hepatic tissues, may lead to bleeding (esophagus) varices which is lethal if not treated urgently.

In a previous study [5], we performed a Newtonian flow analysis for the portal vein, and also simulated the hepatic circulation variations in after a virtual hepatectomy. This paper can be viewed as a sister-paper of [5] in that, we intended to investigate whether Newtonian blood flow model is a good approximation of PV flow, or if a non-Newtonian model should be employed instead. The decision may be derived from a numerical analysis of the shear strain rate, which was shown in Fig. 5. The figure indicates that low shear rate regions ($5s^{-1}$) were presented in a small portion of wall area under normal inflow conditions and thus a Newtonian model could be sufficient. This is verified by a comparison between a non-Newtonian model (the Carreau model) and the Newtonian model. The comparisons of the flow velocity profile and pressure gradient on an identical cross-section suggested that there were only marginal differences between the two models, although the flow profile yielded from the non-Newtonian model is slightly more flat and even. This may be explained from the fact that the higher viscosity simulated from the Carreau model has a higher resistance to flow, or a larger ‘dissipative’ force to distribute blood flow [6].

However, it should be emphasized that under pathological conditions, flow rate in the PV could be substantially lower, and thus the flow simulation can be very different with a non-Newtonian model. This was confirmed in Fig. 8(b) when the portal flow was only 10% of normal value. This might have physiological significance when it comes to biomedical research since the Newtonian model underestimated the WSS. Our future work, therefore, is to collect experimental data on shear rate and shear stress in the portal vein.

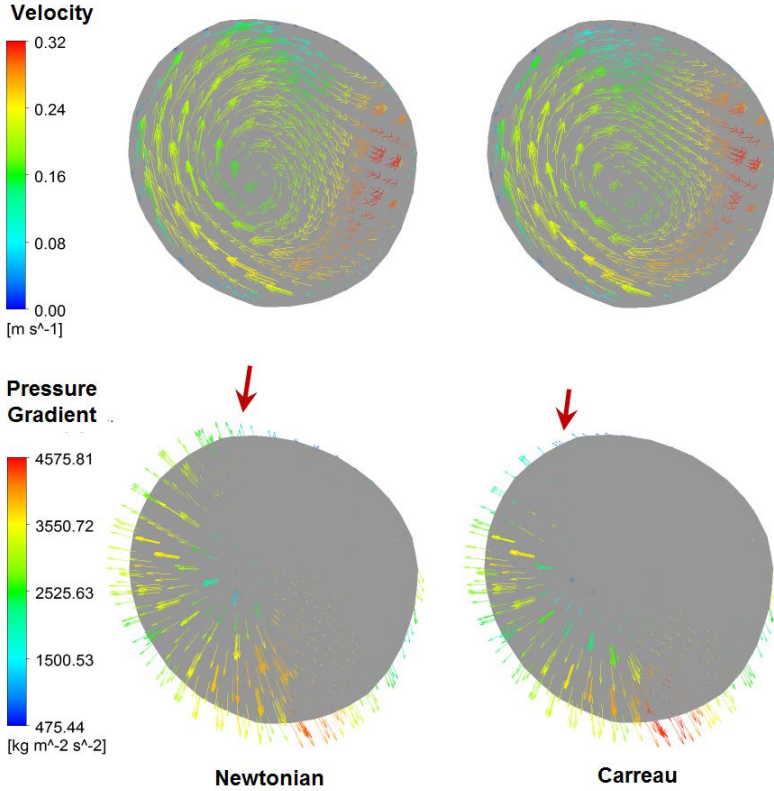


Fig. 7. Comparison of (a) flow velocity and (b) pressure-gradient on the cross-section. The differences between the two models are marginal.

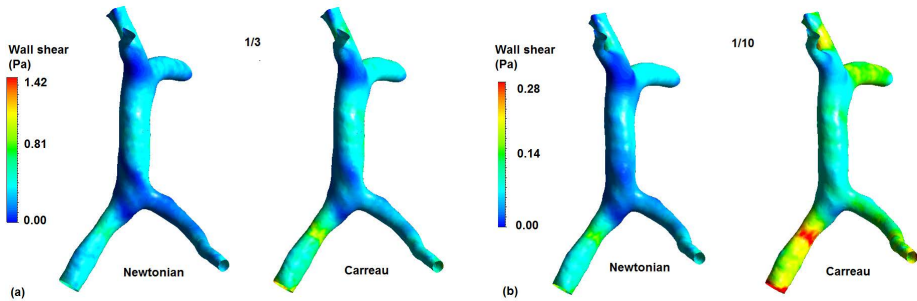


Fig. 8. Comparison of Newtonian and non-Newtonian models under (a) 1/3 of normal flow rate; and (b) 1/10 of normal flow rate. The differences between the two models are significant.

5 Conclusion

We simulated blood flow in a patient-specific portal vein which was digitized from a CT image. We found that under normal physiological conditions there were only marginal difference between the simulation results yielded from the Newtonian and non-Newtonian models. However, in pathological conditions non-Newtonian effects became significant due to substantially lower flow rate in the portal vein.

Acknowledgements. This work was financially supported by the Vice Chancellor Strategical Research Fund of the University of Auckland, which we gratefully acknowledge.

References

1. Frydrychowicz, A., Landgraf, B.R., Niespodzany, E., Verma, R.W., Roldán-Alzate, A., Johnson, K.M., Wieben, O., Reeder, S.B.: Four-dimensional velocity mapping of the hepatic and splanchnic vasculature with radial sampling at 3 tesla: A feasibility study in portal hypertension. *Journal of Magnetic Resonance Imaging* 34(3), 577–584 (2011)
2. Fung, Y.C.: *Biomechanics: Mechanical Properties of Living Tissues*, 2nd edn. Springer-Verlag New York, Inc., New York (1993)
3. George, S.: *Hemodynamic Investigation of the Liver using Magnetic Resonance Imaging and Computational Fluid Dynamics*. PhD thesis, Georgia Institute of Technology (2008)
4. Ho, C.-M., Lin, R.-K., Tsai, S.-F., Hu, R.-H., Liang, P.-C., Wen-Hann Sheu, T., Lee, P.-H.: Simulation of portal hemodynamic changes in a donor after right hepatectomy. *Journal of Biomechanical Engineering* 132(4), 041002–041007 (2010)
5. Ho, H., Bartlett, A., Hunter, P.: Blood Flow Simulation for the Liver after a Virtual Right Lobe Hepatectomy. In: *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2012*. LNCS, vol. 7510, pp. 513–520. Springer, Heidelberg (2012)
6. Ho, H., Wu, J., Hunter, P.: Blood flow simulation in a giant intracranial aneurysm and its validation by digital subtraction angiography. In: Wittek, A., M.F. Nielsen, P., Miller, K. (eds.) *Computational Biomechanics for Medicine*, pp. 15–26. Springer, Heidelberg (2011)
7. Johnston, B.M., Johnston, P.R., Corney, S., Kilpatrick, D.: Non-Newtonian blood flow in human right coronary arteries: steady state simulations. *Journal of Biomechanics* 37(5), 709–720 (2004)
8. Merrill, E.W., Gilliland, E.R., Cokelet, G., Shin, H., Britten, A., Wells Jr., R.E.: Rheology of human blood, near and at zero flow: Effects of temperature and hematocrit level. *Biophysical Journal* 3(3), 199–213 (1963)
9. Niemann, C.U., Roberts, J.P., Ascher, N.L., Yost, C.S.: Intraoperative hemodynamics and liver function in adult-to-adult living liver donors. *Liver Transplantation* 8(12), 1126–1132 (2002)

Tracer Kinetic Modeling by Morales-Smith Hypothesis in Hepatic Perfusion CT

Sang Ho Lee, Wenli Cai, and Hiroyuki Yoshida

3D Imaging Research, Department of Radiology,
Massachusetts General Hospital and Harvard Medical School,
25 New Chardon St., Suite 400C, Boston, Massachusetts 02114, USA
{lee.sangho, cai.wenli, yoshida.hiro}@mgh.harvard.edu

Abstract. Most of the existing tracer kinetic models for dynamic contrast-enhanced CT or MRI do not fully describe the principles of intra- and transcapillary transport of tracers. One point is to disregard the concentration profiles between the inlets and outlets of capillaries, which may cause a biased estimation of tissue parameters by a systematic error. The Morales-Smith hypothesis enables one to resolve this ambiguity by assuming that the difference between arterial and venous concentrations is proportional to the difference between the arterial and capillary concentrations. If the backflow of administered tracer into the plasma compartment is negligible compared to its outflow into the interstitial compartment during the initial enhancement phase after tracer administration, the capillary concentration can be considered to fall exponentially along the capillary from the arterial concentration to the venous concentration by the Renkin-Crone model, i.e., unidirectional extraction fraction, which can be incorporated in the concept of the Morales-Smith hypothesis. In this study, we reformed the mass-balance equations and mathematical solutions of several representative and well-known tracer kinetic models so that the Morales-Smith hypothesis could be incorporated into their compartment tracer kinetics, considering a tissue-specific factor independent of time as proposed by Brix et al. [5]. The tissue-specific factor was applied to a liver tumor case study in perfusion CT to illustrate the potential effectiveness of the Morales-Smith hypothesis. The proposed scheme was shown to be potentially useful for more consistent and reliable estimation of physiologic tissue parameters.

Keywords: Tracer kinetic modeling, Morales-Smith hypothesis, Renkin-Crone model, hepatic perfusion CT.

1 Introduction

Dynamic contrast-enhanced (DCE) imaging with CT or MRI has received widespread interest because these techniques allow noninvasive in vivo assessment of tissue hemodynamics and could provide physiologic information about the tissue microvasculature [1-3].

Assessment of hemodynamic changes in the liver is particularly challenging because of the dual blood supply to this organ [3]. Various degrees of abnormalities in

the rate of blood supply from the hepatic artery and portal vein can be found according to the different kinds of liver diseases [1, 4]. Thus, a reliable estimate of a total hepatic perfusion to consider a tissue-specific factor [5] as well as separate measurements of hepatic arterial and portal venous perfusions are particularly important for the liver diseases.

The microvascular characteristics of diseased tissues can be investigated by fitting of a tracer kinetic model on a region of interest or pixel basis to tracer concentration time course data derived from DCE imaging data. A number of different tracer kinetic models, with varying degrees of complexity, have been used for quantification of the vascular properties of normal and tumor tissues [3, 5-8].

The question arises which tracer kinetic model is optimal for quantification of DCE data, under what conditions, and how large the differences in physiologic parameters are. For selecting or justifying a model of a physiologic system, the different assumptions related to concepts of tracer kinetics must be considered, starting with a simplified model which is successfully expanded, keeping in mind that there is always a trade-off between model complexity and estimability. In addition, if a particular model selected cannot be qualified as an absolute best kinetic model, as circumstances require, selecting statistically the physiologic parameters provided for good discrimination between non-cancerous and cancerous tissues across different kinetic models would be needed [9].

Thus, we first review several representative and well-known pharmacokinetic models, and we compare their kinetic parameter maps in hepatic perfusion CT. For further study, we incorporated the Morales-Smith hypothesis [10] to all compartment tracer kinetic models presented in order to resolve the ambiguity of capillary concentration from the arterial concentration to the venous concentration by considering a tissue-specific factor independent of time, as proposed by Brix et al. [5]. Our primary aim in this study was to illustrate the effectiveness of the Morales-Smith hypothesis through comparison between kinetic parameter maps applied to a liver tumor case study in perfusion CT.

2 Methods

We consider two sources of blood flow to the liver: flow from the hepatic artery, F_A , and flow from the portal vein, F_{PV} . Assuming that the tracer plasma concentrations for the hepatic artery $C_A(t)$ (g/min) and the portal vein $C_{PV}(t)$ can be sampled from dynamic CT images, the concentration for the liver tissue $C_T(t)$ can be expressed as follows:

$$\begin{aligned} C_T(t) &= R(t) \otimes \left[\frac{\tilde{F}_A}{V_T} C_A(t) + \frac{\tilde{F}_{PV}}{V_T} C_{PV}(t) \right] \\ &= \frac{\tilde{F}}{V_T} R(t) \otimes [\gamma C_A(t) + (1 - \gamma) C_{PV}(t)], \end{aligned} \quad (1)$$

where $R(t)$, \tilde{F}_A (ml/min), \tilde{F}_{PV} , \tilde{F} , γ and V_T (ml) denote the impulse residue response function, apparent hepatic arterial flow, apparent portal venous flow, apparent total hepatic plasma flow, arterial fraction, and tissue volume, respectively. Thus, $\frac{F_A}{V_T}$ (ml/min/ml), $\frac{F_{PV}}{V_T}$ and $\frac{F}{V_T}$ become the apparent arterial perfusion, apparent portal perfusion, and total apparent hepatic perfusion, respectively. The apparent plasma flow, $\tilde{F} = rF$, is systemically higher than the true plasma flow F by a factor of $1 \leq r \leq 2$, where r is given by the following expression [5]:

$$r = \frac{E}{1 - E \left(\frac{F}{PS} \right)} \text{ with } E = 1 - e^{-\frac{PS}{F}}. \tag{2}$$

Note that E and PS (ml/min) are the extraction fraction given by the Renkin-Crone model [11, 12], and the permeability surface area product, respectively. The factor r is a tissue-specific model parameter that is independent of time, which can be used to approximate the true plasma flow F from the apparent plasma flow \tilde{F} caused by the difference between the capillary plasma and venous concentrations.

For pharmacokinetic analysis of hepatic perfusion CT data, five pharmacokinetic models were formulated: 1) the two compartment exchange (2CX) model, 2) adiabatic approximation to the tissue homogeneity (AATH) model, 3) distributed (DP) parameter model, 4) Tofts-Kety (TK) model, and 5) extended Tofts-Kety (ETK) model. A graphical representation of the five models is given in Fig. 1.

2.1 Two Compartment Exchange (2CX) Model

The 2CX model [5] assumes two compartments that are composed of the plasma and interstitial spaces. These compartments are assumed to be well mixed, i.e., tracer

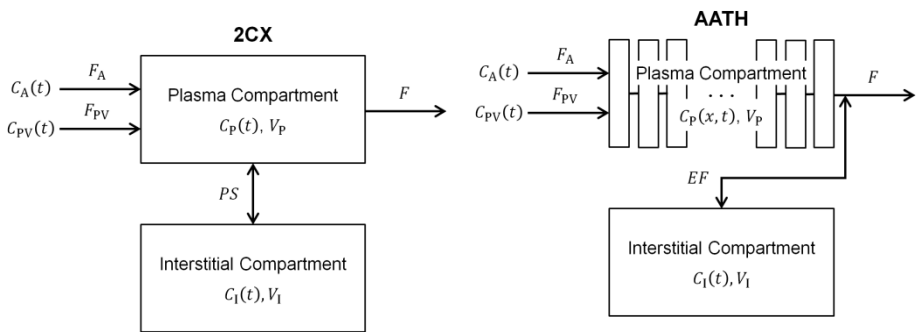


Fig. 1. A graphical representation of the two compartment exchange model (2CX), adiabatic approximation to the tissue homogeneity (AATH) model, distributed parameter (DP) model, Tofts-Kety (TK) model, and extended Tofts-Kety (ETK) model. Note that the TK ($F \ll PS$) and ETK ($F \gg PS$) models are reduced versions of the 2CX model under specific physiological scenarios.

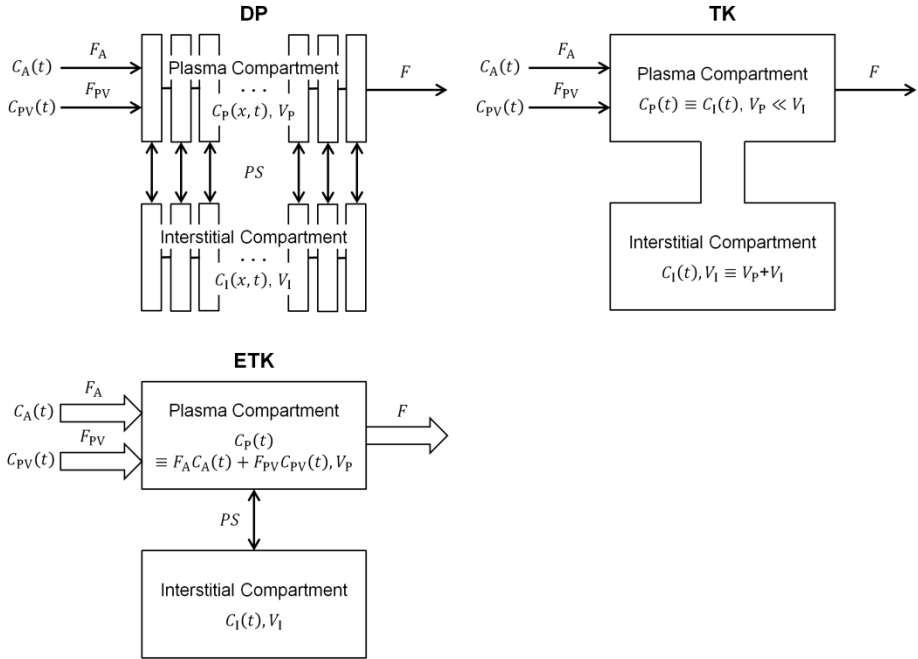


Fig. 1. (continued)

movement is sufficiently fast and distributes evenly throughout the compartment, so that the tracer concentration is a function only of time, but not of space. Thus, the 2CX model is called a lumped-parameter model, which can be described by the following mass-balance equations:

$$\frac{dC_P(t)}{dt} = \frac{\tilde{F}}{V_P} [\gamma C_A(t) + (1 - \gamma) C_{PV}(t) - C_P(t)] + \frac{PS}{V_P} [C_P(t) - C_I(t)], \quad (3)$$

$$\frac{dC_I(t)}{dt} = \frac{PS}{V_I} [C_P(t) - C_I(t)]. \quad (4)$$

Note that $V_P = v_P V_T$ and $V_I = v_I V_T$ are the plasma and interstitial volumes, where v_P and v_I denote the plasma and interstitial volume fractions, respectively. By convolution with the impulse residue function $R(t)$, the analytic solution for $C_T(t)$ can be given as

$$C_T(t) = v_P \frac{\tilde{F}}{V_P} \left[\frac{1}{\alpha - \beta} \left(\left(\alpha + \left(1 + \frac{v_P}{v_I} \right) \frac{PS}{V_P} \right) e^{\alpha t} + \left(\beta + \left(1 + \frac{v_P}{v_I} \right) \frac{PS}{V_P} \right) e^{\beta t} \right) \right] \otimes [\gamma C_A(t) + (1 - \gamma) C_{PV}(t)], \quad (5)$$

where $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{1}{2} \left[- \left(\frac{\tilde{F}}{V_P} + \left(1 + \frac{v_P}{v_I} \right) \frac{PS}{V_P} \right) \pm \sqrt{\left(\frac{\tilde{F}}{V_P} + \left(1 + \frac{v_P}{v_I} \right) \frac{PS}{V_P} \right)^2 - 4 \frac{v_P}{v_I} \frac{\tilde{F}}{V_P} \frac{PS}{V_P}} \right]$.

2.2 Adiabatic Approximation to the Tissue Homogeneity (AATH) Model

The AATH model [8] assumes a plug-flow model for the plasma and a well-mixed compartment for the interstitial space, where the plasma concentration is described by a one-dimensional axial variation $C_P(x, t)$, assuming a capillary tube with a fluid flow. Thus, if the axial length of a capillary tube is L , the average concentration at the venous end is $C_P(L, t)$. In particular, the capillary walls are impermeable in the AATH model, and instead the interstitial space receives influx with clearance EF from the venous end of the capillary. The rest fraction $(1 - E)F$ and the outflux from the interstitial space drain straight to the vein. Because the capillary walls are impermeable in the AATH model, the mass-balance equation in the capillary bed becomes only the plug-flow equation without decay. Because $C_P(L, t)$ is the concentration at the venous end of the capillary, the equation for the interstitial space is described with influx $EF C_P(L, t)$. Additionally in this study, we defined an apparent extraction fraction $\tilde{E} = 1 - e^{-\frac{PS}{\tilde{F}}}$ in order to apply the Morales-Smith hypothesis, i.e., the apparent plasma flow \tilde{F} and the permeability surface area product PS are used for calculation of \tilde{E} . Although the AATH model assumes impermeable capillary walls, this approach can be reasonable because E is related mathematically to PS by the Renkin-Crone model. Therefore, these assumptions lead to the following set of mass-balance equations:

$$\frac{\partial C_P(x, t)}{\partial t} = \frac{\tilde{F}}{V_P} \left[(\gamma C_A(t) + (1 - \gamma) C_{PV}(t)) \delta(x) - L \frac{\partial C_P(x, t)}{\partial x} \right], \tag{6}$$

$$\frac{dC_I(t)}{dt} = \frac{\tilde{E}\tilde{F}}{V_I} [C_P(L, t) - C_I(t)], \tag{7}$$

where $\delta(x)$ is the Dirac delta function that denotes the idealized impulse excitation of a unit-mass source. The analytic solution for $C_T(t)$ becomes

$$C_T(t) = v_P \frac{\tilde{F}}{V_P} \left[u(t) + \left(\tilde{E} e^{-\frac{v_P \tilde{E} \tilde{F}}{v_I V_P} \left(t - \frac{V_P}{\tilde{F}} \right)} - 1 \right) u \left(t - \frac{V_P}{\tilde{F}} \right) \right] \otimes [\gamma C_A(t) + (1 - \gamma) C_{PV}(t)], \tag{8}$$

where $u(t)$ denotes the unit step function.

2.3 Distributed Parameter (DP) Model

The DP model [3] describes the capillary bed as a plug-flow system like the AATH model. However, unlike the AATH model, the interstitial compartment is modeled as

a series of infinitesimal compartments that exchange tracer only with nearby locations in the capillary bed. Thus, the concentrations $C_P(x, t)$ and $C_I(x, t)$ both rely on the position of a capillary tube. The DP model does not allow for axial tracer transport in the interstitial space. Therefore, a tracer cannot travel to the venous ends of the capillary bed through the interstitial space. Thus, the DP model can be considered as a chain of infinitesimal 2CX models. The mass-balance equations in the DP model can be formulated for an elemental volume dx along the axial length L of a capillary tube as follows:

$$\frac{\partial C_P(x, t)}{\partial t} = \frac{\tilde{F}}{V_P} \left[(\gamma C_A(t) + (1 - \gamma) C_{PV}(t)) \delta(x) - L \frac{\partial C_P(x, t)}{\partial x} \right] + \frac{PS}{V_P} [C_P(x, t) - C_I(x, t)], \tag{9}$$

$$\frac{\partial C_I(x, t)}{\partial t} = \frac{PS}{V_I} [C_P(x, t) - C_I(x, t)], \tag{10}$$

The analytical solution for $C_T(t)$ can be expressed as

$$C_T(t) = v_p \frac{\tilde{F}}{V_P} \cdot \left[\left(u(t) + e^{-\frac{PS}{\tilde{F}}} u \left(t - \frac{V_P}{\tilde{F}} \right) \right) \left(1 + \frac{PS}{V_P} \int_0^{t - \frac{V_P}{\tilde{F}}} e^{-\frac{v_P PS}{v_I V_P \tau}} \sqrt{\frac{v_P V_P}{v_I \tilde{F}} \frac{1}{\tau}} I_1 \left(2 \frac{PS}{V_P} \sqrt{\frac{v_P V_P}{v_I \tilde{F}} \tau} \right) d\tau \right) \right] \otimes [\gamma C_A(t) + (1 - \gamma) C_{PV}(t)], \tag{11}$$

where I_1 denotes the modified Bessel function of the first kind.

2.4 Tofts-Kety (TK) and Extended Tofts-Kety (ETK) Models

Under certain conditions such as limited information on the data or the tracer administration, the use of the 2CX model may be inadequate for data analysis. In these cases, a reduced model can be used for the identification of reliable tissue parameters. The ETK model [6] is a reduced version of the 2CX model, which holds for a scenario in which the plasma flow is sufficiently high to replenish loss of tracer into the interstitial space, i.e. a permeability-limited model ($PS \ll F$). In terms of an operation, the ETK model assumes that the concentration time curve in the plasma compartment cannot be distinguished from the arterial input concentration, i.e., $C_P(t) \equiv \gamma C_A(t) + (1 - \gamma) C_{PV}(t)$ for the liver. The $C_T(t)$ for the ETK model is obtained from the equation (5) for the 2CX model in the limit $F \rightarrow \infty$ as

$$C_T(t) = v_P[\gamma C_A(t) + (1 - \gamma)C_{PV}(t)] + v_P \frac{PS}{V_P} e^{-\frac{v_P PS}{v_1 V_P} t} \otimes [\gamma C_A(t) + (1 - \gamma)C_{PV}(t)]. \quad (12)$$

Although the ETK model is originally the permeability-limited model, it can be converted into a mixed flow- and permeability-limited or clearance model by use of the relationship $PS \equiv \tilde{E}\tilde{F}$. Thus, equation (12) can be reformulated as

$$C_T(t) = v_P[\gamma C_A(t) + (1 - \gamma)C_{PV}(t)] + v_P \frac{\tilde{E}\tilde{F}}{V_P} e^{-\frac{v_P \tilde{E}\tilde{F}}{v_1 V_P} t} \otimes [\gamma C_A(t) + (1 - \gamma)C_{PV}(t)]. \quad (13)$$

Thus, this conversion enables one to apply the Morales-Smith hypothesis, as can be done for the other models presented. Note that $r = 2$ as $F \rightarrow \infty$ in the ETK model, i.e., $\tilde{F} = 2F$. On the other hand, the TK model [7] ignores an intravascular tracer, i.e., $C_T(t) = v_1 C_1(t)$, which can be considered in terms of the flow-limited condition ($PS \gg F$) in the 2CX model. The flow-limited model is obtained from equation (5) for the 2CX model in the limit $PS \rightarrow \infty$ as

$$C_T(t) = v_P \frac{\tilde{F}}{V_P} e^{-\frac{v_P \tilde{F}}{(v_P + v_1) V_P} t} \otimes [\gamma C_A(t) + (1 - \gamma)C_{PV}(t)]. \quad (14)$$

Because $v_P \ll v_1$ in the TK model, $\frac{v_P}{v_P + v_1} \equiv \frac{v_P}{v_1}$. In addition, under the flow-limited condition, $\tilde{F} \equiv \tilde{E}\tilde{F}$. Thus, equation (14) can be reformulated as

$$C_T(t) = v_P \frac{\tilde{E}\tilde{F}}{V_P} e^{-\frac{v_P \tilde{E}\tilde{F}}{v_1 V_P} t} \otimes [\gamma C_A(t) + (1 - \gamma)C_{PV}(t)]. \quad (15)$$

Note that $r = 1$ as $PS \rightarrow \infty$ in the TK model, i.e., $\tilde{F} = F$ and $\tilde{E} = E$.

3 Results

We used a hepatocellular carcinoma (HCC) case to generate kinetic parameter maps from different pharmacokinetic models and to compare them between with and without the Morales-Smith hypothesis in perfusion CT. Imaging was performed on a 16-MDCT scanner (LightSpeed, GE Healthcare). Four consecutive slices in the region of the HCC were selected for cine image acquisition. A dynamic contrast enhanced study of the selected four slices was performed in a single breath-hold at the end of expiration with a static table position. As the contrast-media protocol for this case, a total of 70 ml of nonionic iodinated contrast medium (300 mg/ml) was injected in the antecubital vein at the rate of 7 ml/s through an 18-gauge IV cannula by use of a power injector. We used the following CT parameters to obtain dynamic data:

gantry rotation time, 1 s; 100 kVp; 240 mA; acquisition in four images per gantry rotation; and a reconstructed slice thickness of 5 mm. Six seconds after the start of contrast medium injection, an initial scan was obtained, followed by continuous image acquisition with a temporal resolution of 0.5 s for a total duration of 25 s, depending on the patient’s maximum breath hold capacity. Intermittent imaging was then performed once every 15 s for a total scan duration of 225 s.

Fig. 2 shows the parameter maps for the apparent hepatic perfusion ($\frac{\tilde{F}}{V_T}$), apparent arterial perfusion ($\frac{\tilde{F}_A}{V_T}$), and apparent portal perfusion ($\frac{\tilde{F}_{PV}}{V_T}$), which are calculated by

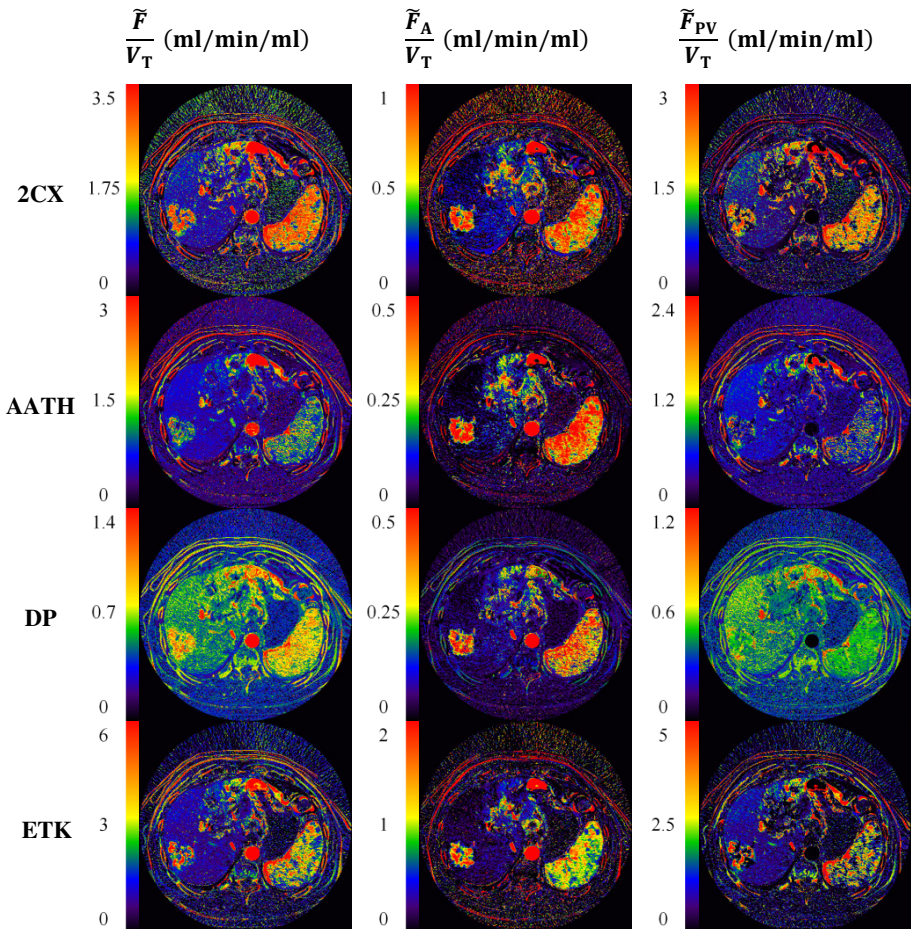


Fig. 2. Parameter maps for apparent hepatic perfusion ($\frac{\tilde{F}}{V_T}$), apparent arterial perfusion ($\frac{\tilde{F}_A}{V_T}$), and apparent portal perfusion ($\frac{\tilde{F}_{PV}}{V_T}$), calculated with a constant factor $r = 1$

assuming a constant factor $r = 1$ without the Morales-Smith hypothesis. In Fig. 3, the parameter maps for the hepatic perfusion ($\frac{F}{V_T}$), arterial perfusion ($\frac{F_A}{V_T}$), and portal perfusion ($\frac{F_{PV}}{V_T}$) are corrected by the variable tissue-specific factor r for applying the

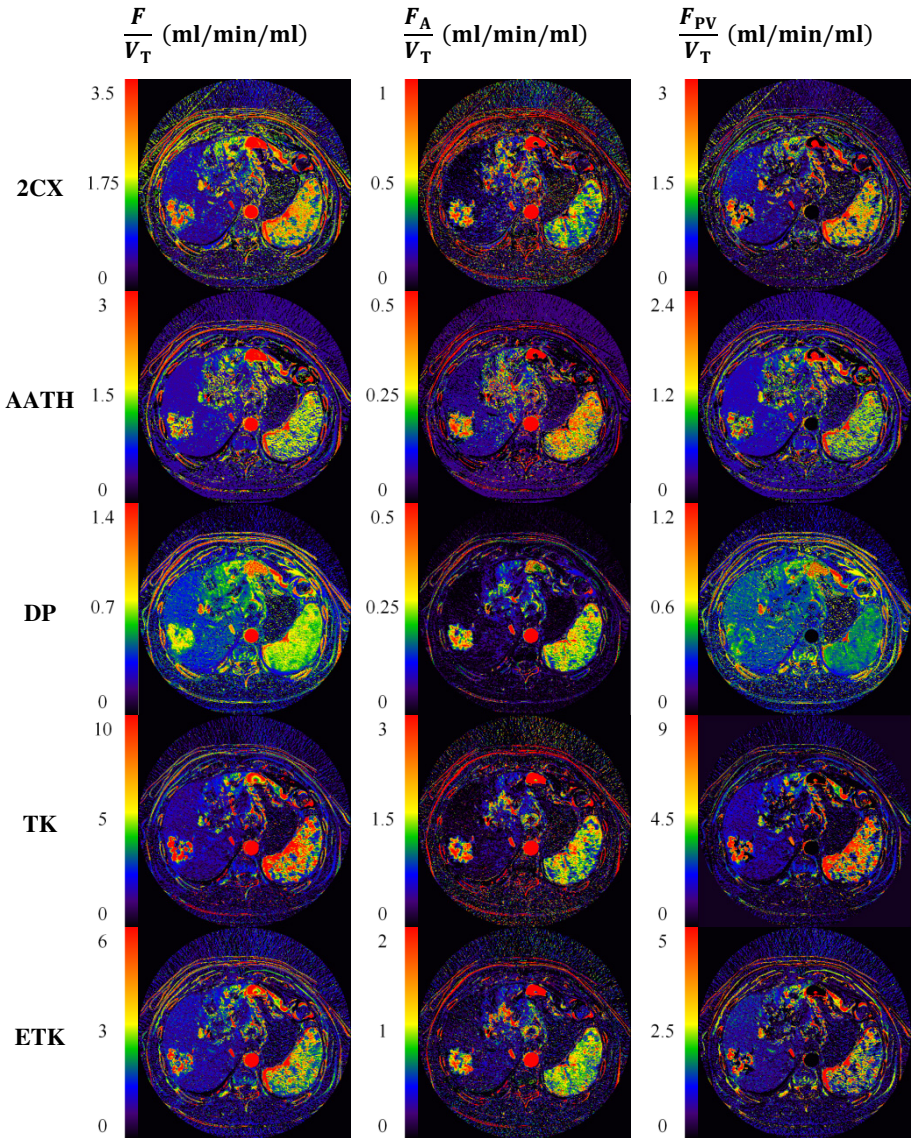


Fig. 3. Parameter maps for hepatic perfusion ($\frac{F}{V_T}$), arterial perfusion ($\frac{F_A}{V_T}$), and portal perfusion ($\frac{F_{PV}}{V_T}$), corrected by the Morales-Smith hypothesis

Morales-Smith hypothesis. The apparent arterial perfusion was relatively overestimated in the tumor tissue as compared to the corrected arterial perfusion, whereas the apparent portal perfusion exhibited relatively uneven mapping patterns in both normal and tumor tissues as compared to the corrected portal perfusion, affecting their total apparent hepatic perfusion. When the apparent hepatic perfusion with the corrected hepatic perfusion was compared among different kinetic models, the variation in the apparent hepatic perfusion was relatively higher in both normal and tumor tissues, as theoretically expected because the venous concentration hidden at the outlet of capillaries, which cannot be measured noninvasively, can influence the capillary plasma flow. In addition, the range of the apparent and corrected perfusion values mapped was different among different kinetic models because of their different underlying physiologic assumptions. Overall, the corrected hepatic perfusion, corrected arterial perfusion, and corrected portal perfusion improved the mapping quality, the consistency of the perfusion values within the same type of tissue, and the contrast between normal and tumor tissues. Note that the tissue-specific factor r is not applied to the TK model because of the constant value of $r = 1$. The ETK model was relatively less affected by the tissue-specific factor r because of the constant value of $r = 2$.

4 Conclusion

The preliminary results of this study showed that the model solution considering a tissue-specific factor by the Morales-Smith hypothesis provides promising prospects for more consistent quantification of well-defined physiologic parameters that characterize the tissue microcirculation and microvasculature.

References

1. Pandharipande, P.V., Krinsky, G.A., Rusinek, H., Lee, V.S.: Perfusion Imaging of the Liver: Current Challenge and Future Goals. *Radiology* 234, 661–673 (2005)
2. Sahani, D.V., Holalkere, N.S., Mueller, P.R., Zhu, A.X.: Advanced Hepatocellular Carcinoma: CT Perfusion of Liver and Tumor Tissue - Initial Experience. *Radiology* 243, 736–743 (2007)
3. Koh, T.S., Thng, C.H., Lee, P.S., Hartono, S., Rumpel, H., Goh, B.C., Bisdas, S.: Hepatic Metastases: In Vivo Assessment of Perfusion Parameters at Dynamic Contrast-Enhanced MR Imaging with Dual-Input Two-Compartment Tracer Kinetics Model. *Radiology* 249, 307–320 (2008)
4. Van Beers, B.E., Leconte, I., Materne, R., Smith, A.M., Jamart, J., Horsmans, Y.: Hepatic Perfusion Parameters in Chronic Liver Disease: Dynamic CT Measurements Correlated with Disease Severity. *Am. J. Roentgenol.* 176, 667–673 (2001)
5. Brix, G., Zwick, S., Kiessling, F., Griebel, J.: Pharmacokinetic Analysis of Tissue Microcirculation Using Nested Models: Multimodel Inference and Parameter Identifiability. *Med. Phys.* 36, 2923–2933 (2009)
6. Tofts, P.S.: Modeling Tracer Kinetics in Dynamic Gd-DTPA MR Imaging. *J. Magn. Reson. Imaging* 7, 91–101 (1997)

7. Tofts, P.S., Brix, G., Buckley, D.L., Evelhoch, J.L., Henderson, E., Knopp, M.V., Larsson, H.B., Lee, T.Y., Mayr, N.A., Parker, G.J., Port, R.E., Taylor, J., Weisskoff, R.M.: Estimating Kinetic Parameters from Dynamic Contrast-Enhanced T(1)-Weighted MRI of a Diffusible Tracer: Standardized Quantities and Symbols. *J. Magn. Reson. Imaging* 10, 223–232 (1999)
8. St. Lawrence, K.S., Lee, T.Y.: An Adiabatic Approximation to the Tissue Homogeneity Model for Water Exchange in the Brain: I. Theoretical Derivation. *J. Cereb. Blood Flow Metab.* 18, 1365–1377 (1998)
9. Lee, S., Kim, J., Cho, J., Kim, S., Song, I., Kim, H., Kim, S.: Multifarious Kinetic Analysis for Differentiation of Prostate Cancer and Benign Prostatic Hyperplasia in DCE-MRI. In: *Proc. Intl. Soc. Mag. Reson. Med.*, vol. 19, p. 2639 (2011)
10. Morales, M., Smith, R.E.: On the Theory of Blood-Tissue Exchange of Inert Gases: VI. Validity of Approximate Uptake Expressions. *Bull. Math. Biophys.* 10, 191–200 (1948)
11. Renkin, E.M.: Transport of Potassium-42 from Blood to Tissue in Isolated Mammalian Skeletal Muscles. *Am. J. Physiol.* 197, 1205–1210 (1959)
12. Crone, C.: The Permeability of Capillaries in Various Organs as Determined by Use of the 'Indicator Diffusion' Method. *Acta. Physiol. Scand.* 58, 292–305 (1963)

Author Index

- Azencott, Robert 158
- Bartels, Lambertus W. 229
- Bartlett, Adam 283
- Boone, Darren J. 1, 10
- Buhmann, Joachim M. 97
- Cai, Wenli 88, 292
- Cattin, Philippe C. 137
- Chang, Stephen 178
- Chen, Wenyu 246
- Chen, Xinjian 208
- Chui, Chee Kong 178
- d'Alché-Buc, Florence 274
- de Groen, Piet C. 116
- Fujiwara, Michitaka 238
- Fundana, Ketut 137
- Gil, Debora 265
- Gladish M.D., Gregory 158
- González, Miguel A. 265
- Gu, Xianfeng David 20
- Halligan, Steve 1, 10
- Hampshire, Thomas E. 1, 10
- Hawkes, David J. 1, 10
- Ho, Harvey 283
- Hostettler, Alexandre 198
- Hu, Mingxing 10
- Huang, Weimin 178, 246
- Hunter, Peter 283
- Imuta, Masanori 40
- Irie, Toshiyuki 255
- Jain, Saurabh 158
- Jin, Hyeong-min 58
- Kakadiaris, Ioannis A. 158
- Kim, Chang Won 49
- Kim, Jong Hyo 49, 58
- Kim, Se Hyung 30
- Kirchhoff, Sonja 107
- Kiriyanthan, Silja 137
- Kitasaka, Takayuki 238
- Klink, Camiel 168
- Kobatake, Hidefumi 126
- Kurihara, Tsuneya 255
- Lee, June-Goo 88
- Lee, Sang Ho 78, 292
- Letort, Véronique 274
- Li, Zisheng 255
- Liang, Jerome Zhengrong 20
- Lim, Calvin 178
- Linguraru, Marius George 188, 265
- Liu, Jianfei 188
- Liu, Jiang 178
- Lucidarme, Olivier 274
- Mahapatra, Dwarikanath 97
- Majeed, Tahir 137
- Manh Luu, Ha 168
- Marescaux, Jacques 198
- Matsuzaki, Kazuki 255
- Matuszewski, Bogdan J. 218
- McClelland, Jamie R. 10
- Mendrik, Adrienne 168
- Misawa, Kazunari 238
- Moelker, Adriaan 168
- Moore, Christopher 218
- Mori, Kensaku 238
- Muthukudage, Jayantha 116
- Näppi, Janne J. 30, 40, 69, 78
- Navab, Nassir 107
- Nawano, Shigeru 126
- Nawarathna, Ruwan 116
- Nicolau, Stephane 198
- Niessen, Wiro 168
- Noorda, Yolanda H. 229
- Oda, Masahiro 238
- Oh, JungHwan 116
- Oost, Elco 126
- Papadakis, Manos 158
- Papiez, Bartłomiej W. 218

- Piel, Christina 88
 Plum, Josien P.W. 229
 Punwani, Shonit 10
- Regge, Daniele 69
 Reiser, Maximilian F. 107
 Rémond, Yves 198
 Rockey, Don 69
 Romain, Blandine 274
 Roth, Holger R. 1, 10
 Rouet, Laurence 274
- Samei, Golnoosh 147
 Schueffler, Peter 97
 Shark, Lik-Kwan 218
 Shi, Rui 20
 Shimizu, Akinobu 126
 Slabaugh, Greg 1
 Soler, Luc 198
 Su, Yi 178
 Summers, Ronald M. 188
 Székely, Gabor 147
- Tanner, Christine 147
 Tavanapong, Wallapak 116
 Taylor, Stuart 10
 Tian, Qi 246
- Tielbeek, Jeroen A.W. 97
 Tomoshige, Sho 126
- Upadhyay, Sanat 158
- van Walsum, Theo 168
 Venkatesh, Sudhakar K. 246
 Vera, Sergio 265
 Vos, Franciscus M. 97
- Wang, Shijun 188
 Watanabe, Hidefumi 126
 Wong, Johnny 116
- Xiong, Wei 178, 246
- Yamashita, Yasuyuki 40
 Yang, Tao 178
 Yang, Zepa 58
 Yao, Jianhua 208
 Yigitsoy, Mehmet 107
 Yoshida, Hiroyuki 30, 40, 69, 78, 88, 292
- Zeng, Wei 20
 Zhang, Da 88
 Zhang, Jing 178
 Zhao, Hemin 208
 Zhou, Jiayin 178, 246
 Zhu, Wenwu 198