Hiroyuki Yoshida
Simon Warfield
Michael W. Vannier (Eds.)

LNCS 8198

Abdominal Imaging

Computation and Clinical Applications

5th International Workshop
Held in Conjunction with MICCAI 2013
Nagoya, Japan, September 2013, Proceedings

Lecture Notes in Computer Science 8198

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Hiroyuki Yoshida Simon Warfield
Michael W. Vannier (Eds.)

Abdominal Imaging

Computation and Clinical Applications

5th International Workshop
Held in Conjunction with MICCAI 2013
Nagoya, Japan, September 22, 2013
Proceedings

Springer

Volume Editors

Hiroyuki Yoshida
Massachusetts General Hospital and Harvard Medical School
3D Imaging Research, Department of Radiology
Boston, MA 02114, USA
E-mail: yoshida.hiro@mgh.harvard.edu

Simon Warfield
Boston Children's Hospital and Harvard Medical School
Department of Radiology
Boston, MA 02115, USA
E-mail: simon.warfield@childrens.harvard.edu

Michael W. Vannier
The University of Chicago
Department of Radiology
Chicago, IL 60637, USA
E-mail: mvannier@radiology.bsd.uchicago.edu

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-41082-6 e-ISBN 978-3-642-41083-3
DOI 10.1007/978-3-642-41083-3
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013948468

CR Subject Classification (1998): I.4.9, I.4.1, I.4.3, I.4.6-9, J.3, I.2.10, F.1.1

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

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 Fifth International Workshop on Abdominal Imaging: Computational and Clinical Applications, was held in conjunction with the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) on September 22, 2013, in Nagoya, Japan.

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 fifth 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 38 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 many cases, by 3 expert reviewers from the Scientific Review Committee. Based on the results of this review, 32 papers were accepted by the workshop for presentation. 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 further provided two plenary lectures, one on diffusion-weighted MRI analysis of Crohn's disease in the bowel by Dr. Simon Warfield from Boston Children's Hospital and Harvard Medical School; the other on applications of perfusion CT/MRI to liver and pancreatic diseases by Dr. Yoshihisa Tsujii from Kyoto University Hospital, in order to introduce participants' current clinical challenges that were not fully explored in the MICCAI main conference. As a result, 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, and the members of the Editorial Board for their outstanding job in compiling the papers in this proceedings volume.

September 2013

Hiroyuki Yoshida
Simon Warfield
Michael W. Vannier

Workshop Organization

Organizing Committee

Hiroyuki Yoshida	Massachusetts General Hospital/Harvard Medical School, USA
Simon Warfield	Boston Children's Hospital/Harvard Medical School
Michael Vannier	The University of Chicago, USA

Program Chair

Janne Näppi	Massachusetts General Hospital/Harvard Medical School, USA

Local Organizer

Masahiro Oda	Nagoya University, Japan

Workshop Committee

Philippe C. Cattin	Medical Image Analysis Center/University of Basel, Switzerland
Aaron Fenster	Robarts Research Institute/University of Western Ontario, Canada
Hiroshi Fujita	Gifu University, Japan
Lubomir Hadjiyski	University of Michigan, USA
Kenneth R. Hoffman	University at Buffalo, USA
David Hawkes	University College London, UK
David R. Holmes	Mayo Clinic College of Medicine, USA
Lakhmi C. Jain	University of South Australia, Australia
Jong Hyo Kim	Seoul National University, South Korea
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
Daniele Regge	Institute for Cancer Research and Treatment, Italy
Akinobu Shimizu	Tokyo University of Agriculture and Technology, Japan

Yoshihisa Shinagawa	Siemens, USA
Frans Voss	Delft University of Technology/Academic Medical Center Amsterdam, The Netherlands

Program Committee

Dongqing Chen	GE Healthcare, USA
Yufei Chen	Tongji University, China
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
Moti Freiman	Children's Hospital Boston/Harvard Medical School, USA
Masahiro Oda	Nagoya University, Japan
Danielle Pace	Kitware Inc., USA
Frank Preiswerk	University of Basel c/o University Hospital Basel, Switzerland
Holger Roth	University College London, UK
Sergio Vera	Universitat Autònoma de Barcelona, Spain
Xin Zhao	State University of New York at Stony Brook, USA
Hongbin Zhu	State University of New York at Stony Brook, USA

Scientific Review Committee

Philippe C. Cattin	Medical Image Analysis Center/University of Basel, Switzerland
Dongqing Chen	GE Healthcare, USA
Yufei Chen	Tongji University, China
Hiroshi Fujita	Gifu University, Japan
Lubomir Hadjiyski	University of Michigan, USA
Kenneth R. Hoffman	University at Buffalo, USA
David R. Holmes	Mayo Clinic College of Medicine, USA
David Hawkes	University College London, UK
Jong Hyo Kim	Seoul National University, South Korea
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
Moti Freiman	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
Danielle Pace	Kitware Inc., USA
Frank Preiswerk	University of Basel c/o University Hospital Basel, Switzerland
Holger Roth	University College London, UK
Yoshinobu Sato	Osaka University Graduate School of Medicine, Japan
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
Michael Vannier	The University of Chicago, USA
Sergio Vera	Universitat Autònoma de Barcelona, Spain
Hiroyuki Yoshida	Massachusetts General Hospital/Harvard Medical School, USA
Xin Zhao	State University of New York at Stony Brook, USA

Proceedings Editorial Board

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

Colon and Other Gastrointestinal Tract - Virtual Colonoscopy

Liver, Kidney, and Other Organs

A Model Development Pipeline for Crohn's Disease Severity Assessment from Magnetic Resonance Images

Peter J. Schüffler[1,*], Dwarikanath Mahapatra[1], Jeroen A.W. Tielbeek[2], Franciscus M. Vos[2, 3], Jesica Makanyanga[4], Doug A. Pendsé[4], C. Yung Nio[2], Jaap Stoker[2], Stuart A. Taylor[4, 5], and Joachim M. Buhmann[1]

[1] Dept of Computer Science, ETH Zurich, Universitätstrasse 6, Zurich, CH
[2] Dept of Radiology, Academic Medical Center, Meibergdreef 9, Amsterdam, NL
[3] Quantitative Imaging Group, Delft University of Technology, Lorentzweg 1, Delft, NL
[4] Centre for Medical Imaging, University College London, 250 Euston Road, London, UK
[5] Dept of Radiology, University College Hospital London, 235 Euston Road, London, UK
peter.schueffler@inf.ethz.ch

Abstract. Crohn's Disease affects the intestinal tract of a patient and can have varying severity which influences treatment strategy. The clinical severity score CDEIS (Crohn's Disease Endoscopic Index of severity) ranges from 0 to 44 and is measured by endoscopy. In this paper we investigate the potential of non-invasive magnetic resonance imaging to assess this severity, together with the underlying question which features are most relevant for this estimation task. We propose a new general and modular pipeline that uses machine learning techniques to quantify disease severity from MR images and show its value on Crohn's Disease severity assessment on 30 patients scored by 4 medical experts. With the pipeline, we can obtain a magnetic resonance imaging score which outperforms two existing reference scores MaRIA and AIS.

Keywords: Crohn's Disease, abdominal MRI, CDEIS, MaRIA, AIS.

1 Introduction

Crohn's Disease is a chronic Inflammatory Bowel Disease that often affects the terminal ileum and colon causing inflammation, stenoses, fistula and ulcers. Symptoms of the disease include abdominal pain, diarrhea and weight loss due to a malfunction of the bowel. While the exact cause of Crohn's Disease (CD) is not known, it is thought to be a multifactorial mixture of environmental influences having an adverse effect on the immune system of genetically predisposed people. The treatment of patients in different stages of the disease include autoimmune suppressives and antibiotics as well as surgery of affected parts of the bowel in severe cases. CD patients undergo a regularly examination in which the severity of CD is determined. The severity and activity of the disease has direct influence on the current treatment strategy.

* Corresponding author.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 1–10, 2013.

A state-of-the-art score for this disease severity is the Crohn's Disease Endoscopic Index of Severity (CDEIS), which is determined by ileo-colonoscopy. Severity scores are assigned to the five bowel segments *rectum, sigmoid and descend colon, transverse colon, ascended colon, terminal ileum.* Three points are added to the mean segmental score if ulcerated stenoses or non-ulcerated stenoses are present to define the patient's CDEIS. Colonoscopy in general is time consuming, uncomfortable (partly painful) for the patient and has inherent limitations as e.g. hindered accessibility after stenoses. Therefore, there is ongoing research into alternative imaging methods such as magnetic resonance imaging (MRI). In this paper, we investigate to what extent MRI can serve as basis for the measurement of CD severity. We make following contributions: 1) developing a new systematic pipeline for model generation for CD severity assessment based on MRI; 2) proposing a plausible method for feature selection within this pipeline; 3) developing a new MRI based severity score for Crohn's Disease with the help of this pipeline. The new score is evaluated with its Pearson correlation to the CDEIS and compared to existing scores. Our findings can help in determining the relevant features to be addressed when it comes to automated MRI analysis in this context. Fully automated localization, calibration and segmentation are recently central bottlenecks for computer driven feature extraction.

1.1 Related Work

Two published MRI based CD related scores are used as reference: the MaRIA model [1] and the AIS [2]. The MaRIA score serves as a baseline for our experiments, since it is optimized for CDEIS correlation. However, it uses the critical feature of relative contrast enhancement (RCE), which is very time consuming and highly subjective to measure. The scores are defined as:

$$\boldsymbol{MaRIA} = 1.5 * wall\ thickness + 0.02 * RCE + 5 * edema + 10 * ulceration$$

$$\boldsymbol{AIS} = 1.79 + 1.34 * mural\ thickness + 0.94 * mural\ T2\ signal$$

2 Methods

To develop a MRI based severity score, we propose a new systematic machine learning pipeline, which in principal can be applied on similar problems in computational radiology (Fig. 1). (1) Driven by the target score CDEIS which is calculated on each bowel segment, the first step in the pipeline is segment-wise feature extraction from MRI volumes. (2) An exhaustive search is systematically performed throughout all features for feature and model selection. This step incorporates combinatorial feature selection, which are subjected to linear regression models, and patient-wise cross-validation of the models. After ranking the models according to the median correlation to CDEIS, we discover a class of models with similar performance. (3) We propose a method to identify this set by P-values and show how a feature distribution in this set can support the model selection in step 3. (4) The selected model is validated on a separated dataset which has not been used in steps 1-3.

Fig. 1. Model development pipeline proposed in this paper. Each step is modular and can be adjusted to various kinds of biomedical problems. The feature distribution in this pipeline is an unconventional but very effective step.

2.1 Dataset and Feature Extraction

In total, we use the data of 30 CD patients, of which 20 patients were used for brute force model generation and ranking (including cross-validation) and 10 patients for independent testing. These numbers are competitive for this disease (e.g. 16 patients for training / 26 for testing in [2]).

MRI scans of 30 luminal CD patients with written consent of data usage were collected at the Academic Medical Center (AMC), Amsterdam [3]. For bowel distention, patients drank 1,6l of Mannitol (2.5%, Baxter, Utrecht, The Netherlands). 1 hour later, MRI sequences were aquired with a 3.0T scanner (Intera, Philips Healthcare, Best, The Netherlands): axial and coronal T2-weighted single shot fast spin echo (SSFSE) sequences w/ and w/o fat saturation, as well as a coronal 3D T1-weighted spoiled gradient echo (SPGE) sequence with fat saturation. Thereafter, the injected contrast agent butylscopalaminebromide (20mg, Buscopan, Boehringer, Ingelheim, Germany) helped for a dynamic contrast enhanced (DCE-MRI) MRI sequence with gadobutrol (0.1 ml/kg, Gadovist 1.0 mmol/ml, Bayer Schering Pharma, Berlin, Germany). A coronal DCE-MRI sequence with 450 scans over 6 min (temporal resolution: 0.82s, spatial resolution: 2.78x2.78x2.5mm) was performed. Then, a second dose of butylscopalaminebromide (20mg) was injected, followed by postcontrast axial and coronal 3D T1-weighted SPGE sequences with fat saturation [3]. The DCE-MRI sequence was not used in this study. Fig. 2 shows two typical images of two patients. The scans were independently visually examined by four radiology experts. Each patient's bowel was partitioned into five segments. Bowel segments are visually identified by the radiologists in a standardized way: e.g. the ileocoecal valve, splenic flexure and hepatic flexure were used to separate the different colon segments. Each segment is then visually rated according to 17 CD related features by every expert:

Per bowel segment: *mural_thickness, muralT2, perimural_T2, pattern, enhancement_T1, comb_sign, abcess, fistula, length, wall_thickness, rce, edema, ulcers and pseudopolyps,*
Per patient: *lymph_nodes_pP, node_enhancement_pP and enlarged_lymphnodes_pP*

These features comprise MRI features described in the literature and used by most abdominal radiologists as identified in an international inventory, as well as those used in the two published scoring systems. For a medical description of these features,

please see [4]. Note that there is no computer screened feature available for CD, yet. We focus exactly on this aspect by investigating which visible features are relevant. However, research on automated MRI analysis with automated feature extraction is ongoing [5, 6].

Out of the potential 150 bowel segments, 7 segments could not be assessed by the radiologists and 6 segments were not accessible by endoscopy, resulting in 137 usable bowel segments with features and label. Every bowel segment has one CDEIS score as label from the colonoscopy as described before (score ranges from 0 to 38), allocated by one expert gastroenterologist. Since most segments are not affected by CD, 96 segments have a CDEIS score 0 and 41 segments a score greater than 0.

The inter-observer variance among the MRI features between the 4 radiologists is relatively low, implying that the feature findings in the scans are reproducible (*cf.* [4]). To cover the low variance between different domain experts and to exclude a potential bias according to one expert, the samples of the four observers are bundled together resulting in a dataset of 137*4=548 samples. Each sample has 17 features. 188 samples of 10 random patients were separated for later independent testing.

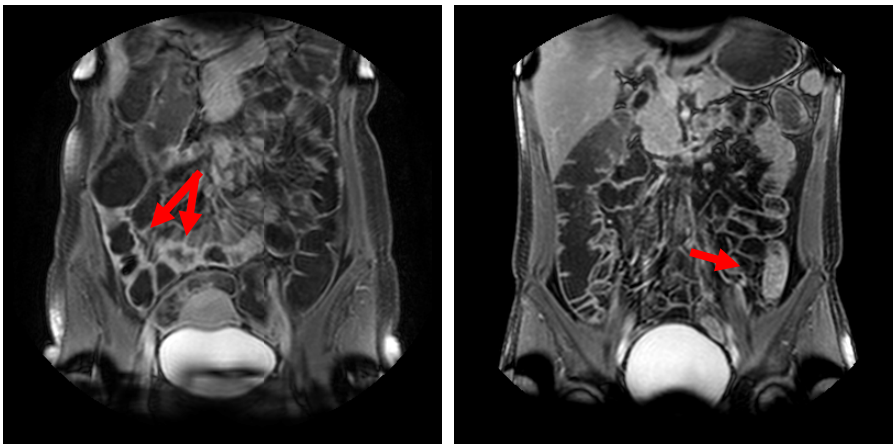


Fig. 2. Two typical MR images from two patients, showing parts of the small bowel and colon. Enhanced regions were annotated by a MD to be associated with CD.

2.2 Feature / Model Selection

An exhaustive search algorithm throughout all 17 features was conducted for feature selection. All 131071 possible combinations of the features were subjected to linear regression models [7], which are easily interpretable, computationally fast, less prone to overtraining and easily comparable to other approaches [8], [1] and [2]. Each model was validated in a 50-fold patient wise bootstrapped cross-validation experiment on the data of our 20 patient data (360 samples). In this scenario, 20 patients are randomly drawn *with replacement* out of the available 20 patients, resulting in roughly

13 different patients per draw. The model is trained on the samples of these patients and tested on the remaining 7 out-of-bag patients. The tested Pearson correlation is recorded and the procedure is repeated 50 times. In every iteration, a different random sample is drawn from the 20-patients dataset. Due to the highly variable nature of biological data, and due to the limited amount of training data, this bootstrapped validation approach provides a realistic simulation to mimic the heterogeneity of the whole population. Note that for every model, the same bootstrap-folds are drawn in each iteration to facilitate an accurate comparison of the different models. The stratification into complete patients is necessary to account for potential dependencies among samples within one patient. All data of a specific patient is contained either in the training set or in the test set, but not distributed over both sets during one iteration.

3 Results

3.1 Model for Severity Assessment

In the exhaustive search approach, all 131071 (2^{17}-1) possible linear regression models were ranked by their median cross-validated correlation to the targeted CDEIS. The top-ranked "**Model 1**" has a cross-correlation of r=.65, which tends to be higher than the MaRIA (r=.56, P=0.07, Paired t-test) and is significantly higher than the AIS (r=.55, P=0.01), even if they are retrained (both r=.60, P<0.01), meaning an improvement of 8-18 %.

$$\boldsymbol{Model1} = 2.89 * enhancement\ T1 + 7.08 * combsign + 4.95 * edema + 16.62 * ulcers$$

Model 1 uses only 4 of 17 features, illustrating that more features do not necessarily improve the correlation. Fig. 3 shows the cross-validated correlation of model 1 to CDEIS, compared to two alternative scores (Model 23 and 63), a random model, the MaRIA and the AIS. We will explain Model 23 and 63 in the next section.

3.2 Feature/Model Distribution

Although the top model shows a significant high performance, it may not be reported as the final problem solution, since the second best and third best models (and so on) show similar results in the cross-validation (Fig. 3). Actually, the correlation coefficients of the best 116 models do not differ significantly from Model 1 (P<0.05, Bonferroni corrected for multiple testing [9]).

Fig. 4A shows the ranked median correlations of all 131071 possible models with a color code for the number of features used by the corresponding models. From a machine learning point of view, it is difficult to justify the top ranked model to be the final solution, while there exists a class of models with statistically similar performance. We solve this problem with a feature distribution within this class, which

enables selecting favorite models according to additional criteria. As depicted in Fig. 4B, the feature distribution among the best 116 models reveals the prominent importance of the features *comb_sign*, *ulcers* and *enhancement_T1*, as they appear in nearly all of these models. The model using only these three features is ranked on position 23 and has a median correlation to CDEIS of r=.64.

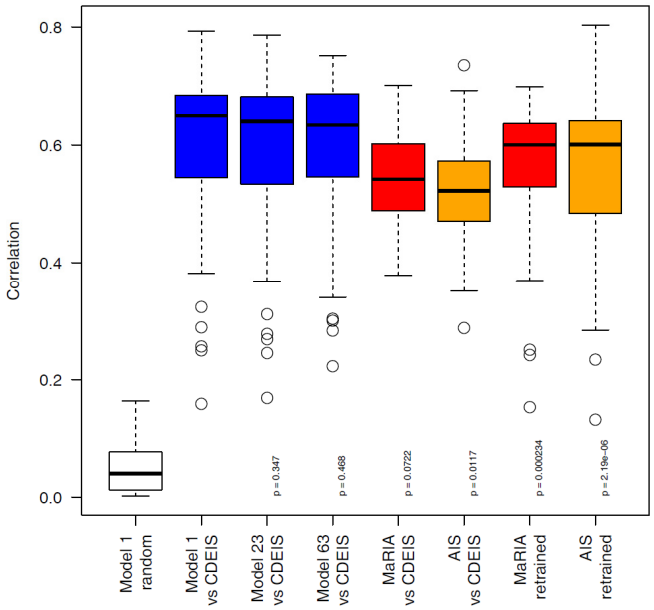


Fig. 3. Pearson correlation performance of our top 3 MRI models. Y-axis: absolute correlation coefficient. For comparison, a random model (left, white) is trained on the features of Model 1, but with randomly permuted CDEIS label before cross-validation. Top ranked models 1, 23 and 63 (blue) have a median correlation of r=.65 to r=.63. The MaRIA (red) and AIS (orange) are cross-validated on our dataset with and without retraining (i.e. adjusting the weights to our dataset and using the weights as proposed by the models (see Introduction)). P-values (paired t-test) express the significant differences to Model 1.

$$\boldsymbol{Model23} = 4.51 * enhancement\ T1 + 8.21 * combsign + 18.37 * ulcers$$

If a low number of features is wished, we can deliver Model 23 as a proper solution of the problem. Indeed, the contribution of additional features to the three will not result in statistically higher cross-validated correlation. Nevertheless, we continue with Model 1, as a low number of features is not mandatory for the specific CD problem.

Another criterion for model selection might be low variance during cross-validation. Model 63 (r=.63) shows lowest variance among the top ranked models:

$$\boldsymbol{Model63} = 2.06 * muralT2 + 7.19 * combsign + 3.12 * length + 0.03 * rce$$

As in our specific problem, it is more important not to use the expensive RCE feature than to have low variance, we do not consider Model 63 as final solution.

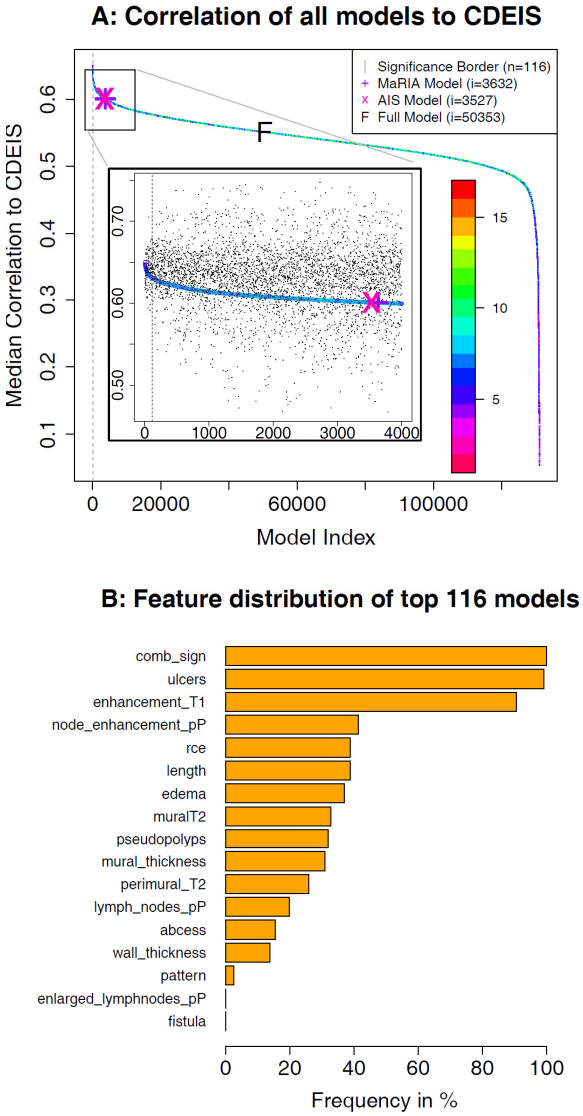


Fig. 4. A: Median cross-validated Pearson correlations of all 131071 possible models. X-axis: model index of ranked models. Color encodes the number of features used per model. The univariate models have weaker correlation. Models 1-116 (gray vertical line in zoom rectangle) do not differ statistically significantly to model 1 (paired t-test $p \geq 0.05$, Bonferroni corrected for multiple testing). +, x and F show the position of the models using MaRIA features, AIS features or all 17 features, respectively. Additional to the colored median CV correlation, the zoom rectangle shows to every model the corresponding correlation on the external test set. **B:** Feature distribution of MRI features in the best 116 models, which do not differ significantly in performance (gray line in Fig. 4A). The top three of the predictors were used in nearly 100% of all 116 models, illustrating their importance for regression.

3.3 Independent Model Testing

Since the result of the cross-validation experiment was used for model selection (by ranking), an independent validation step is needed to consider the issue of overfitting. This test performance should ideally perform in the same range as the cross-validation correlation. As stated in section Methods, we use the samples of 10 randomly chosen patients of the original data as test set. All other 20 patients are then used for training. The test results are reported in Table 1. Model 1 outperforms all other models on the test set by 6-25%. The significance of the differences of the models in the test set are tested in a leave-one-out cross-correlation.

Table 1. Pearson correlation to CDEIS computed by bootstrapped cross-validation and on 10 independent patients. Although MaRIA and AIS are fully parameterized models, we considered them either with **ret**rained weights (ret) or **app**lied without retrained weights (app) before testing. The independent test set is stratified by the four observers. N is the number of bowel segments in the sets. P-values of paired t-tests show the partly significant differences to Model 1.

		Model 1	MaRIA ret	AIS ret	MaRIA app	AIS app
Cross-Validation (n=360)	μ ± σ median t-test P	.59 ± .15 .65	.56 ± .12 .60 2.3e-4	.56 ± .14 .60 2.2e-6	.54 ± .07 .54 .072	.52 ± .09 .52 .012
Test set (n=188)	μ ± σ median t-test P	**.69 ± .08** **.71**	.57 ± .06 .57 6.2e-318	.64 ± .15 .65 7.3e-19	.59 ± .07 .60 6.2e-256	.67 ± .13 .67 1.0e-166

4 Discussion and Conclusion

Four radiologists exhaustively recorded observations concerning 17 features per bowel segment in 30 patients. We could train classifiers to predict a score with high correlation to the real CDEIS, with r=0.65 as highest correlation factor. Compared to the already published MaRIA score, this is an improvement of 18% on our dataset. Note that the MaRIA score was reported to have a much better performance (r=0.82, [8]). This difference might result from the MaRIA score being originally developed on data with more severe CD cases and higher CDEIS, which tend to have clearer

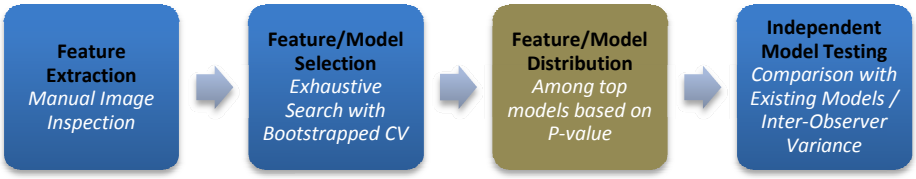


Fig. 5. The systematic model development pipeline as implemented in this paper. Applicable on various feature selection problems, we demonstrate the workflow on Crohn's disease data.

occurrences of the different features in the MRI scans. Also, the MaRIA patients experienced colonic preparation which results in better mucosal visualization at MRI. These effects might be due to a regional variation.

Our proposed feature selection workflow can principally be applied on various feature selection problems. For the CD problem, we implemented the development pipeline with following specifications (Fig. 5): In step 1, we use manually collected features by four domain experts. The ultimate vision of computational MRI processing might require automated feature extraction. However, to the best of our knowledge, no automated feature extraction for computational radiology tailored to CD severity assessment does exist. Our work reveals the importance of single features and, therefore, identifies targets that can be addressed for automated processing in further studies. In step 2, an exhaustive search among all 17 features has revealed that there is no clear prominent model solely suitable for CDEIS representation. Rather, we can have a set of models with different feature combinations but similar performance. This effect can at best been seen using a brute force algorithm instead of a heuristic approach, since in the first, really all models are considered. Further, this observation can often be made in analogue biomedical problems (e.g. [10]). The reasons for that might be manifold. First, there might be no evidence that the features do contain such a prominent relation to the examined target variable. Second, candidate features might indeed be related to the target, but not necessarily causally (directly), such that a mathematical model would only discover indirect (or "weaker") relations. These indirect relations are then discovered in an exhaustive search approach. Finally, from a machine learning point of view, implicit noise in real data, and subjectively measured features and labels always exacerbate the analyses of relations between the measurements. The pipeline respects this observation in step 3: based on the P-value, the set of top performing models is identified. Further, the feature distribution in this set supports the model selection process. Features that occur more often among the top models are assumed to have superior impact on the decision. A final independent validation step is proposed in this pipeline, importantly, since the result of the cross-validation has been used for model selection. Comparing the feature annotations of four different radiologists, our resulting model for CD severity quantification shows superior correlation performance to the state-of-the-art score than existing methods, with at the same time similar variance among different annotators.

Our proposed features and model for CD severity assessment should be further validated on larger datasets for clinical value. An interesting validation scenario would be the connection of MRI with histological data. While histology might serve as a better gold-standard than CDEIS, such data are far less available and often biased to severe cases where surgery was necessary. For future research, our feature ranking offers a basis for automated MRI feature extraction methods which focus on mimicking manual MRI features.

We have presented a systematic pipeline for feature and model selection in computational radiology and showed its benefits for the difficult problem of CD severity estimation based on MRI. With the pipeline, models could be generated with superior performance than existing score functions. We are convinced that this design of a medical imaging pipeline can easily be adapted to structurally similar problems in computational magnetic resonance imaging.

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

References

1. Rimola, J., Ordas, I., Rodriguez, S., Garcia-Bosch, O., Aceituno, M., Llach, J., Ayuso, C., Ricart, E., Panes, J.: Magnetic resonance imaging for evaluation of Crohn's disease: validation of parameters of severity and quantitative index of activity. Inflammatory Bowel Diseases 17, 1759–1768 (2011)
2. Steward, M.J., Punwani, S., Proctor, I., Adjei-Gyamfi, Y., Chatterjee, F., Bloom, S., Novelli, M., Halligan, S., Rodriguez-Justo, M., Taylor, S.A.: Non-perforating small bowel Crohn's disease assessed by MRI enterography: derivation and histopathological validation of an MR-based activity index. European Journal of Radiology 81, 2080–2088 (2012)
3. Ziech, M.L., Lavini, C., Caan, M.W., Nio, C.Y., Stokkers, P.C., Bipat, S., Ponsioen, C.Y., Nederveen, A.J., Stoker, J.: Dynamic contrast-enhanced MRI in patients with luminal Crohn's disease. European Journal of Radiology 81, 3019–3027 (2012)
4. Tielbeek, J.A.W., Makanyanga, J.C., Bipat, S., Pendsé, D.A., Yung Nio, C., Vos, F.M., Taylor, S.A., Stoker, J.: Grading Crohn's disease activity with MRI: Interobserver variability of MRI features, MRI scoring of severity and correlation with Crohn's Disease Endoscopic Index of Severity. AJR (2013)
5. Mahapatra, D., Schueffler, P., Tielbeek, J., Vos, F.M., Buhmann, J.M.: Crohn's Disease Tissue Segmentation from Abdominal MRI Using Semantic Information and Graph Cuts. In: Proc. IEEE ISBI 2013, San Francisco, pp. 358–361 (2013)
6. Mahapatra, D., Schueffler, P., Tielbeek, J.A.W., Buhmann, J.M., Vos, F.M.: A Supervised Learning Based Approach to Detect Crohn's Disease in Abdominal MR Volumes. In: Yoshida, H., Hawkes, D., Vannier, M.W. (eds.) Abdominal Imaging 2012. LNCS, vol. 7601, pp. 97–106. Springer, Heidelberg (2012)
7. Hastie, T., Tibshirani, R., Friedman, J.H.: The elements of statistical learning: data mining, inference, and prediction. Springer, New York (2009)
8. Rimola, J., Rodriguez, S., Garcia-Bosch, O., Ordas, I., Ayala, E., Aceituno, M., Pellise, M., Ayuso, C., Ricart, E., Donoso, L., Panes, J.: Magnetic resonance for assessment of disease activity and severity in ileocolonic Crohn's disease. Gut 58, 1113–1120 (2009)
9. Bonferroni, C.E.: Il calcolo delle assicurazioni su gruppi di teste. In: Studi in Onore del Professore Salvatore Ortu Carboni, Rome, pp. 13–60 (1935)
10. Cima, I., Schiess, R., Wild, P., Kaelin, M., Schueffler, P., Lange, V., Picotti, P., Ossola, R., Templeton, A., Schubert, O., Fuchs, T., Leippold, T., Wyler, S., Zehetner, J., Jochum, W., Buhmann, J., Cerny, T., Moch, H., Gillessen, S., Aebersold, R., Krek, W.: Cancer genetics-guided discovery of serum biomarker signatures for diagnosis and prognosis of prostate cancer. Proceedings of the National Academy of Sciences of the United States of America 108, 3342–3347 (2011)

Spatially Constrained Incoherent Motion (SCIM) Model Improves Quantitative Diffusion-Weighted MRI Analysis of Crohn's Disease Patients

Vahid Taimouri, Moti Freiman, Onur Afacan, and Simon K. Warfield

Computational Radiology Laboratory, Boston Children's Hospital, Harvard Medical School, 300 Longwood Ave, Boston, MA, 02115, USA
vahid.taimouri@childrens.harvard.edu

Abstract. Quantitative analysis of fast and slow diffusion from abdominal Diffusion-weighted MRI has the potential to provide important new insights into physiological and microstructural properties of the body. However, the commonly used, independent voxel-wise fitting of the signal decay model leads to imprecise parameter estimates, which has hampered their practical usage. In this work we evaluated the improvement in the precision of the fast and slow diffusion parameter estimates achieved by using a spatially-constrained Incoherent Motion (SCIM) model of DW-MRI signal decay in 5 healthy subjects and 24 Crohn's disease patients. We found that the improvement in Coefficient of Variation (CV) of the parameter estimates achieved using the SCIM model was significantly larger compared to thus achieved by repeated acquisition and signal averaging (n=5, paired Student's t-test, $p \leq 0.05$). We also found that the SCIM model reduced the coefficient of variation of the parameter estimates of the D^* and f parameter estimates in the ileum by 30% compared to the independent voxel-wise fitting of the signal decay model in the Crohn's patients data (n=24, paired Student's t-test, $p \leq 0.05$). The SCIM model is more precise for quantitative analysis of abdominal DW-MRI signal decay.

Keywords: Diffusion-weighted imaging, Crohn's disease, intra-voxel incoherent motion.

1 Introduction

Diffusion-weighted MRI (DW-MRI) of the body is a non-invasive imaging technique sensitive to the incoherent motion of water molecules inside the area of interest. This motion is known to be a combination of a slow diffusion component associated with the Brownian motion of water molecules, and a fast diffusion component associated with the bulk motion of intravascular molecules in the micro-capillaries. These phenomena are characterized through the so-called, intra-voxel incoherent motion (IVIM) model with the slow diffusion (D); the fast

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 11–19, 2013.

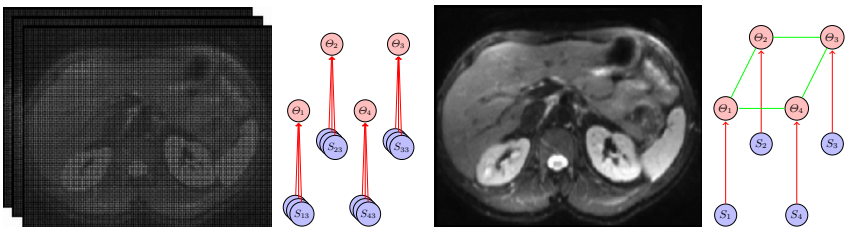


Fig. 1 Illustration of the graphical models used to estimate the fast and slow diffusion parameters from DW-MRI data. (a, b) represents independent voxel-wise estimation of intra-voxel incoherent motion with multiple DW-MRI images averaged, and (c, d) represents voxel-wise estimation of the DW-MRI signal decay model parameters using the spatially constrained incoherent motion model.

diffusion (D^*) as decay rate parameters; and the fractional contribution (f) of each motion to the DW-MRI signal decay [1,2,3].

IVIM model parameters have recently shown promise as quantitative imaging biomarkers for various clinical applications in the body including differential analysis of tumors [4,5,6,7,8,9], the assessment of liver cirrhosis [10,11], and Crohn's disease [12].

However, the utility of IVIM parametric imaging with DW-MRI is diminished by a lack of verified methods for producing reliable estimates of both fast and slow diffusion parameters from the DW-MRI signal [3]. Specifically, reliable estimates of IVIM model parameters are difficult to obtain because of 1) the non-linearity of the IVIM model; 2) the limited number of DW-MRI images as compared to the number of the IVIM model parameters, and; 3) the low signal-to-noise ratio (SNR) observed in body DW-MRI.

In current practice, the reliability of the incoherent motion parameter estimates is increased by acquiring multiple DW-MRI images from the patient; next, average these results, and then use the averaged DW-MRI signal to estimate IVIM model parameters. However, this requires substantially increased acquisition times - an undesirable outcome, especially in children, who generally have difficulty in remaining still for long periods of time [3].

Recently, Freiman et al. introduced a new model of DW-MRI signal decay which utilizes the spatial homogeneity as a constraint in the DW-MRI signal decay [13,14]. Essentially, the Spatially Constrained Incoherent Motion (SCIM) model produces estimates of Incoherent Motion model parameters for all voxels simultaneously, rather than solving for each voxel independently. As a result, the reliability of the incoherent motion parameter estimates from the DW-MRI data is increased without acquiring additional data. Figure. 1 depicts the graphical models for the repeated acquisition method used previously to improve parameter estimates reliability (a,b) compared to the SCIM model (c,d).

In this work, we evaluated the improvement in parameter estimates reliability by means of Coefficient of Variation (CV) achieved by using our SCIM model compared to the utilization of the repeated acquisition and signal averaging technique using abdominal DW-MRI data of 5 healthy volunteers. We also compared the CV of fast and slow diffusion parameter estimates obtained form DW-MRI data of 24 Crohn's disease patients using our SCIM model and the commonly used independent voxel-wise fitting of the IVIM model.

We found that the SCIM model is up to 45% more efficient in improving parameter estimate reliability compared to the repeated acquisition and signal averaging technique (n=5, $p \leq 0.05$). We also found that the SCIM model provides 30% improvement in parameter estimates reliability compared to the independent voxel-wise fitting of the IVIM model in DW-MRI data of Crohn's disease patients ($p \leq 0.05$).

2 Method

2.1 The Intravoxel Incoherent Motion Model

The Intra-Voxel Incoherent Motion (IVIM) model of DW-MRI signal decay assumes a signal decay function of the form [1,2]:

$$m_{v,i} = s_{0,v} \left(f_v \exp(-b_i(D_v^* + D_v)) + (1 - f_v) \exp(-b_i(D_v)) \right) \tag{1}$$

where $m_{i,v}$ is the expected signal of voxel v at b-value=b_i, $s_{0,v}$ is the baseline signal at voxel v; D_v is the slow diffusion decay associated with extravascular water molecules' motion; D_v^* is the fast diffusion decay associated with the intravascular water molecules' motion; and f_v is the fraction between the slow and fast diffusion compartments.

Given the DW-MRI data acquired with multiple b-values, the observed signal (S_v) at each voxel v is a vector of the observed signal at the different b-values: $S_v = \{s_{v,i}\}, i = 1 \ldots N$.

We model the IVIM model parameters at each voxel v as a continuous-valued four-dimensional random variable (i.e. $\Theta_v = \{s_{0,v}, f_v, D_v^*, D_v\}$). Commonly, the IVIM model parameters Θ_v are estimated from the DW-MRI signal S_v using an independent voxel-wise maximum-likelihood estimator:

$$\hat{\Theta_v} = \arg\max_{\Theta_v} p(S_v|\Theta_v) = \prod_{i=1}^{N} p(S_{v,i}|\Theta_v) \tag{2}$$

Using a Gaussian approximation of the non-central χ-distribution of the acquisition noise [15], and taking the negative log of the maximum likelihood estimator; the maximum likelihood estimation takes the form of a least-squares minimization problem:

$$\hat{\Theta_v} = \arg\min_{\Theta_v} \sum_{i=1}^{N} \left(m_{v,i} - s_{v,i}\right)^2 \tag{3}$$

The IVIM model parameters Θ_v are estimated from the DW-MRI signal S_v by solving the least-squares minimization problem (Eq. 3) for each voxel independently using the Levenberg-Marquardt algorithm [16,17]. Initial estimates of the model parameters were obtained with the least squares estimator [18].

2.2 The Spatially Constrained Incoherent Motion (SCIM) Model

Taking the Bayesian perspective, our goal is to find the parametric maps Θ that maximize the posterior probability associated with the maps given the observed signal S and the spatial homogeneity prior knowledge:

$$\hat{\Theta} = \arg\max_{\Theta} p(\Theta|S) \propto p(S|\Theta)p(\Theta) \tag{4}$$

Based on the Hammersley-Clifford theorem [19], by using a spatial prior in the form of a continuous-valued Markov random field, the posterior probability $p(S|\Theta)p(\Theta)$ can be decomposed into the product of node and clique potentials:

$$p(S|\Theta)p(\Theta) \propto \prod_{v} p(S_v|\Theta_v) \prod_{v_p \sim v_q} p(\Theta_{v_p}, \Theta_{v_q}) \tag{5}$$

where $p(\Theta_v|S_v)$ is the data term representing the probability of voxel v to have the DW-MRI signal S_v given the model parameters Θ_v, $v_p \sim v_q$ is the collection of the neighboring voxels according to the employed neighborhood system, and $p(\Theta_{v,p}, \Theta_{v,q})$ is the spatial homogeneity prior in the model.

By taking the negative logarithm of the posterior probability (Eq. 5), the maximum a posteriori (MAP) estimate Θ is equivalent to the minimization of:

$$E(\Theta) = \sum_{v} \phi(S_v; \Theta_v) + \sum_{v_p \sim v_q} \psi(\Theta_{v_p}, \Theta_{v_q}) \tag{6}$$

where $\phi(S_v; \Theta_v)$ and $\psi(\Theta_{v,p}, \Theta_{v,q})$ are the compatibility functions:

$$\phi(S_v; \Theta_v) = -\log p(S_v|\Theta_v), \qquad \psi(\Theta_{v_p}, \Theta_{v_q}) = -\log p(\Theta_{v_p}, \Theta_{v_q}) \tag{7}$$

The data term $\phi(S_v; \Theta_v)$ is given by taking the negative logarithm of the likelihood function, the spatial homogeneity term is defined using the robust L1-norm:

$$\psi(\Theta_{v_p}, \Theta_{v_q}) = \alpha W |\Theta_{v_p} - \Theta_{v_q}| \tag{8}$$

where $\alpha \geq 0$ weights the amount of spatial homogeneity enforced by the model, and W is a diagonal weighting matrix that accounts for the different scales of the parameters in Θ_v.

3 Experimental Results

3.1 Precision of Incoherent Motion Parameter Estimates from In-vivo DW-MRI Data of Healthy Volunteers

We obtained DW-MRI images of 5 health volunteers who underwent research abdominal MRI studies between January 2013 and May 2013. We carried out MR imaging studies of the abdomen using a 1.5-T unit (Magnetom Avanto, Siemens Medical Solutions, Erlangen, Germany) with a body-matrix coil and a spine array coil for signal reception. Free-breathing single-shot echo-planar imaging was performed using the following parameters: repetition time/echo time (TR/TE) = 6800/59 ms; SPAIR fat suppression; matrix size = 192×156; eld of view = 300×260 mm; number of excitations = 1; slice thickness/gap = 5 mm/0.5 mm; 40 axial slices; 8 b-values = 5,50,100,200,270,400,600,800 s/mm^2. A tetrahedral gradient scheme, first proposed in Conturo et al. [20], was used to acquire 4 successive images at each b-value with an overall scan acquisition time of 4 min. Diffusion trace-weighted images at each b-value were generated using geometric averages of the images acquired in each diffusion sensitization direction [21].

We repeated the imaging acquisition six times to get six DW-MRI datasets, each with low Signal to Noise Ratio (DW-MRI$_{\text{low}}$). We averaged the six DW-MRI datasets to achieve high SNR DW-MRI images (DW-MRI$_{\text{high}}$). We estimated the model parameters from: 1) DW-MRI$_{\text{high}}$ using the independent voxel-wise approach (IVIM$_{\text{high}}$); 2) DW-MRI$_{\text{low}}$ using the independent voxel-wise approach (IVIM$_{\text{low}}$), and 3) DW-MRI$_{\text{low}}$ using the Spatially Constrained Incoherent Motion (SCIM) model (SCIM$_{\text{low}}$).

We calculated the precision of the parameter estimates by means of the coefficient of variation (CV) of the parameter estimates at each voxel in the IVIM and SCIM maps of each patient using model-based wild-bootstrap analysis [22,23]. Figure. 2 depicts a representative parametric maps of the upper abdomen. The SCIM model yields smoother, more realistic maps, especially for the f parameter. We set the value of α to 0.01 and the rescaling matrix W diagonal to $\{1.0, 0.001, 0.0001, 0.01\}$ to provide equal weight to each one of the incoherent motion model parameters. Stopping criteria was defined as an energy improvement of less than 0.1 from the initial energy or 500 iterations.

For each subject, we averaged the CV values over the same, three ROIs mentioned above. We first examined the statistical significance of the difference in the precision of the parameter estimates between the IVIM$_{\text{high}}$ and the IVIM$_{\text{low}}$ estimates using a two-tailed paired Student's t-test with $p \leq .05$ as indicating a significant difference. Then, we examined the statistical significance of the difference in the precision of the parameter estimates between the IVIM$_{\text{low}}$ and the SCIM$_{\text{low}}$ estimates using the same test. We performed the statistical analyses with standard statistical software (Matlab® R2010b; The MathWorks, Natick, MA, USA).

Figure. 3 presents the bar-plots of the CV values for each parameter estimates. While the repeated acquisition technique slightly reduced the CV compared to

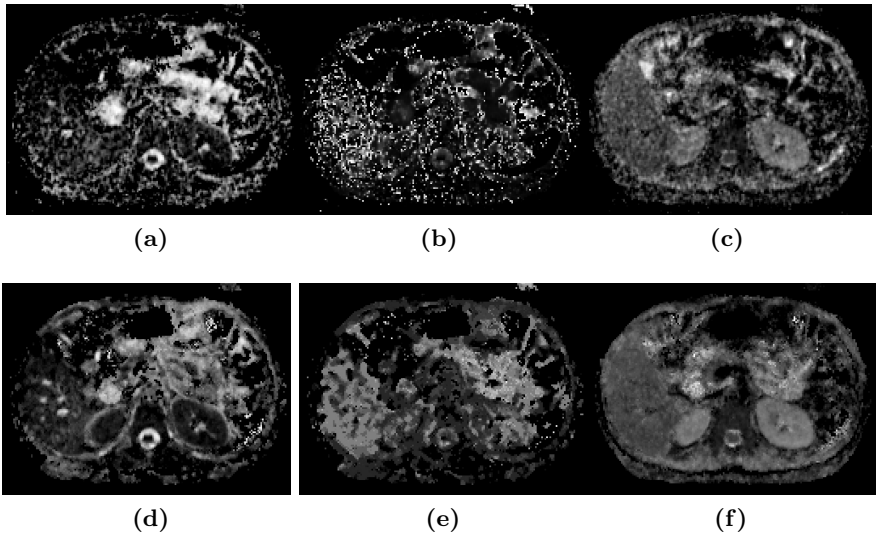


Fig. 2 Representative upper abdomen slice of the parametric maps reconstructed by the IVIM method (1st row), and by the SCIM method (2nd row). The SCIM method yields smoother, more realistic maps, with sensitivity to details.

this of the IVIM_{low}, the difference was not significant. However, by using the SCIM model, the CV of the parameters estimates reduced substantially, by up to 52%. The difference between the CV of the IVIM_{low} and the SCIM_{low} was significant. In addition, we found that the SCIM model is up to 45% more efficient in improving parameter estimate reliability compared to the repeated acquisition and signal averaging technique. As seen in Fig. 3, the error bars of IVIM_{low} are smaller than those of $\text{IVIM}_{\text{high}}$, which implies there is less variation with lower SNR. This issue will be investigated more in the future work.

3.2 Precision of Incoherent Motion Parameter Estimates from In-vivo DW-MRI Data of Crohn's Disease Patients

We acquired DW-MRI and MR enterography (MRE) data from 24 consecutive patients with confirmed Crohn's disease (15 males, 9 females; mean age 14.7 years; range: 5-24 years), who underwent a clinically indicated MRI study between January 1, 2011 and October 31, 2011 in our outpatient MRI department. We carried out MRI imaging studies of the abdomen and the pelvic using a similar protocol to this described in Section 3.1. MR enterography (MRE) protocol for these patients included polyethylene glycol administration for bowel distention and gadolinium-enhanced, dynamic 3D VIBE (volume- interpolated breath hold exam) in the coronal plane.

According to a consensus reading of two board certified radiologists of the MRE data, we classified each patient ileum qualitatively as enhancing or non-enhancing. 11 (46%) patients were diagnosed with abnormal findings in the ileum

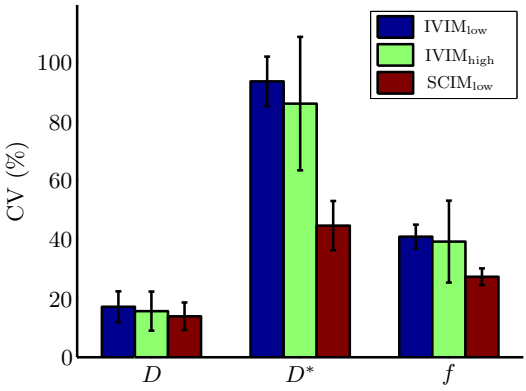
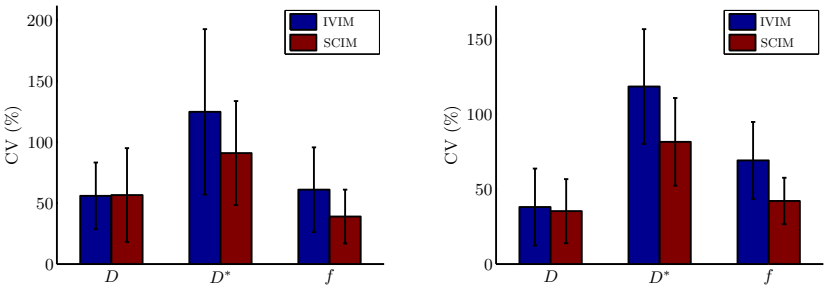


Fig. 3 Bar plot of the CV of the incoherent motion parameters as estimated from 5 healthy subjects. The CV was significantly lower when using our SCIM approach than when using the IVIM approach for all parameters. In contrast, using repeated acquisition and signal averaging did not reduce the coefficient of variation significantly.



(a) Normals **(b)** Abnormals

Fig. 4 Bar plot of the CV of the incoherent motion parameters as estimated from 24 Crohn's disease subjects (13 with normal ileum and 11 with abnormal findings in the ileum). The CV was significantly lower when using our SCIM approach than when using the IVIM approach for all parameters.

and in 13 (54%) patients were diagnosed with normal findings in the ileum. We estimated the signal decay model parameters using the IVIM and SCIM models. We calculated the CV of the parameter estimates for each model. Figure. 4 presents the bar-plots of the CV values for each parameter estimates in the ileum for the normal and abnormal groups. The SCIM model reduced the CV of the f and D^* parameters by 36% and 28% for the normal group, and by 39% and 31% for the abnormal group, respectively, compared to the independent voxel-wise fitting of the signal decay model.

4 Conclusions

The role of incoherent motion parameters as quantitative imaging biomarkers for various clinical applications is becoming increasingly important. However, the current method for estimating the DW-MRI signal decay model are not reliable enough. The reliability of the parameter estimates can be improved substantially by using a spatially constrained model for the signal decay model parameter estimation. In this work, we evaluated the improvement achieved by using our SCIM model using *in-vivo* abdominal DW-MRI data of 5 healthy subjects and 24 Crohn's disease patients. The SCIM model provides a better mechanism to estimate the signal decay model parameters and a more precise insight to the physiological causes of the DW-MRI signal decay and then the voxel-wise independent IVIM model.

References

1. Le Bihan, D.: Intravoxel incoherent motion perfusion MR imaging: a wake-up call. Radiology 249(3), 748–752 (2008)
2. Le Bihan, D., Breton, E., Lallemand, D., Aubin, M.L., Vignaud, J., Laval-Jeantet, M.: Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology 168(2), 497–505 (1988)
3. Koh, D.M., Collins, D.J., Orton, M.R.: Intravoxel incoherent motion in body diffusion-weighted MRI: reality and challenges. AJR Am. J. Roentgenol. 196(6), 1351–1361 (2011)
4. Sigmund, E.E., Cho, G.Y., Kim, S., Finn, M., Moccaldi, M., Jensen, J.H., Sodickson, D.K., Goldberg, J.D., Formenti, S., Moy, L.: Intravoxel incoherent motion imaging of tumor microenvironment in locally advanced breast cancer. Magn. Reson. Med. 65(5), 1437–1447 (2011)
5. Re, T.J., Lemke, A., Klauss, M., Laun, F.B., Simon, D., Grunberg, K., Delorme, S., Grenacher, L., Manfredi, R., Mucelli, R.P., Stieltjes, B.: Enhancing pancreatic adenocarcinoma delineation in diffusion derived intravoxel incoherent motion f-maps through automatic vessel and duct segmentation. Magn. Reson. Med. 66(5), 1327–1332 (2011)
6. Klauss, M., Lemke, A., Grunberg, K., Simon, D., Re, T.J., Wente, M.N., Laun, F.B., Kauczor, H.U., Delorme, S., Grenacher, L., Stieltjes, B.: Intravoxel incoherent motion MRI for the differentiation between mass forming chronic pancreatitis and pancreatic carcinoma. Invest. Radiol. 46(1), 57–63 (2011)
7. Chandarana, H., Lee, V.S., Hecht, E., Taouli, B., Sigmund, E.E.: Comparison of biexponential and monoexponential model of diffusion weighted imaging in evaluation of renal lesions: preliminary experience. Invest. Radiol. 46(5), 285–291 (2011)
8. Gloria, C., Li, Q., Xu, L., Zhang, W.: Differentiation of diffusion coefficients to distinguish malignant and benign tumor. J. Xray Sci. Technol. 18(3), 235–249 (2010)
9. Lemke, A., Laun, F.B., Klauss, M., Re, T.J., Simon, D., Delorme, S., Schad, L.R., Stieltjes, B.: Differentiation of pancreas carcinoma from healthy pancreatic tissue using multiple b-values: comparison of apparent diffusion coefficient and intravoxel incoherent motion derived parameters. Invest. Radiol. 44(12), 769–775 (2009)

10. Luciani, A., Vignaud, A., Cavet, M., Nhieu, J.T., Mallat, A., Ruel, L., Laurent, A., Deux, J.F., Brugieres, P., Rahmouni, A.: Liver cirrhosis: intravoxel incoherent motion MR imaging–pilot study. Radiology 249(3), 891–899 (2008)
11. Patel, J., Sigmund, E.E., Rusinek, H., Oei, M., Babb, J.S., Taouli, B.: Diagnosis of cirrhosis with intravoxel incoherent motion diffusion MRI and dynamic contrast-enhanced MRI alone and in combination: preliminary experience. J. Magn. Reson. Imaging 31(3), 589–600 (2010)
12. Freiman, M., Perez-Rossello, J.M., Callahan, M.J., Bittman, M., Mulkern, R.V., Bousvaros, A., Warfield, S.K.: Characterization of fast and slow diffusion from diffusion-weighted MRI of pediatric Crohn's disease. J. Magn. Reson. Imaging 37(1), 156–163 (2013), http://dx.doi.org/10.1002/jmri.23781
13. Freiman, M., Perez-Rossello, J.M., Callahan, M.J., Voss, S.D., Ecklund, K., Mulkern, R.V., Warfield, S.K.: Reliable estimation of incoherent motion parametric maps from diffusion-weighted MRI using fusion bootstrap moves. Med. Image Anal. 17(3), 325–336 (2013), http://dx.doi.org/10.1016/j.media.2012.12.001
14. Freiman, M., Voss, S.D., Mulkern, R.V., Perez-Rossello, J.M., Callahan, M.J., Warfield, S.K.: Reliable Assessment of Perfusivity and Diffusivity from Diffusion Imaging of the Body. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 1–9. Springer, Heidelberg (2012)
15. Dietrich, O., Raya, J.G., Reeder, S.B., Ingrisch, M., Reiser, M.F., Schoenberg, S.O.: Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics. Magn. Reson. Imaging 26(6), 754–762 (2008)
16. Lemke, A., Stieltjes, B., Schad, L.R., Laun, F.B.: Toward an optimal distribution of b values for intravoxel incoherent motion imaging. Magn. Reson. Imaging 29(6), 766–776 (2011)
17. Yamada, I., Aung, W., Himeno, Y., Nakagawa, T., Shibuya, H.: Diffusion coefficients in abdominal organs and hepatic lesions: evaluation with intravoxel incoherent motion echo-planar MR imaging. Radiology 210(3), 617–623 (1999)
18. Freiman, M., Voss, S.D., Mulkern, R.V., Perez-Rossello, J.M., Callahan, M.J., Warfield, S.K.: In vivo assessment of optimal b-value range for perfusion-insensitive apparent diffusion coefficient imaging. Medical Physics 39, 4832 (2012)
19. Winkler, G.: Image Analysis, Random Fields and Dynamic Monte Carlo Methods, 2nd edn. Springer, New York (2003)
20. Conturo, T.E., McKinstry, R.C., Akbudak, E., Robinson, B.H.: Encoding of anisotropic diffusion with tetrahedral gradients: a general mathematical diffusion formalism and experimental results. Magn. Reson. Med. 35(3), 399–412 (1996)
21. Mulkern, R.V., Vajapeyam, S., Robertson, R.L., Caruso, P.A., Rivkin, M.J., Maier, S.E.: Biexponential apparent diffusion coefficient parametrization in adult vs newborn brain. Magn. Reson. Imaging 19(5), 659–668 (2001)
22. Davidson, R., Flachaire, E.: The wild bootstrap, tamed at last. Journal of Econometrics 146(1), 162–169 (2008)
23. Freiman, M., Voss, S.D., Mulkern, R.V., Perez-Rossello, J.M., Warfield, S.K.: Quantitative Body DW-MRI Biomarkers Uncertainty Estimation Using Unscented Wild-Bootstrap. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part II. LNCS, vol. 6892, pp. 74–81. Springer, Heidelberg (2011)

Self Similarity Image Registration Based on Reorientation of the Hessian

Zhang Li[1], Lucas J. van Vliet[1], and Frans M. Vos[1,2]

[1] Quantitative Imaging Group, Department of Imaging Science and Technology, Delft University of Technology, 2628CJ Delft, The Netherlands
z.li-1@tudelft.nl
[2] Department of Radiology, Academic Medical Center, 1105AZ Amsterdam, The Netherlands

Abstract. The modality independent neighbourhood descriptor (MIND) is a local registration metric that is based on the principle of self-similarity. However, the metric requires recalculation of the self similarity during registration as this inherently changes during image deformation. We propose a self similarity registration method based on the Hessian (HE) that efficiently deals with the recalculation issue. The representation of the local self-similarity via the Hessian enables keeping it up to date during deformation. As such, the registration procedure is efficient and not prone to fall in local minima. We have shown that reorienting the hessian gives a significant improvement ($p<0.05$) over leaving the reorientation out. Our technique also has a better performance over the existing MIND method on the DIR-Lab dataset as well as on abdominal MRI datasets albeit not significant. Ultimately, we will use the technique to quantify Crohn's disease severity based on the relative contrast enhancement in registered images.

Keywords: Image registration, hessian reorientation, Crohn's disease.

1 Introduction

Medical image registration is widely used for finding correspondence between images. Sum of square differences (SSD) and normalized cross covariance (NCC) are often applied for registering images from the same modality. Alternatively, mutual information (MI) is frequently used to deal with multi-modal image registration problems [1–3].

The above, basic similarity metrics are based on global measurements and in principle do not model spatial variance. However, such variance is known influence the robustness of non-rigid registration tasks [4]. Accordingly, Gorbunova et al. [5] proposed a local mass preserving SSD technique for lung CT registration. Likewise, Song et al. [6] used local cross correltation to accommodate inhomogeneities of CT scans. Furthermore, Loeckx et al. [4] proposed a conditional implementation of MI, introducing a spatial dimension into the 3D joint histogram. More approaches combining spatial information into MI can be found in [7–9].

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 20–28, 2013.

It has been noticed [10] that such local estimation can be difficult, though, due to many false local optima in non-rigid registration. An alternative way of incorporating spatial variance was based on structural representations of images. For instance, gradient can be used to find correspondence of images [11, 12]. Moreover, Heinrich et al. [13] proposed a method that relied on orientation information derived from the structure tensor. More recently, the same group [10] introduced the modality independent neighbourhood descriptor (MIND) that is based on the principle of self-similarity. Essentially, this technique assumes that the local image structure is shared by the images to be registered. The method yielded reliable registration results accross different image modalities and better performance than other state-of-the-art approaches. However, the metric requires recalculation of the self similarity during registration as this inherently changes during image deformation.

We propose a self similarity registration method based on the Hessian (HE) that efficiently deals with the recalculation issue. Our method relies on a reorientation strategy adapted from diffusion tensor image registration [14]. A key novelty is that the reorientation of Hessian (ROHE) is integrated in the registration optimization. We compare our procedure with the MIND method on the same dataset from DIR-Lab [15] since MIND is also a self-similarity based method and its performance was previously tested on that dataset. Additionally, we evaluate the technique on abdominal pre- and postcontrast MRI datasets. Ultimately, it is our objective to employ the method to quantify the amount of contrast enhancement in those images, which is known to reflect Crohn's disease severity.

2 Methods

2.1 Modality Independent Neighbourhood Descriptor (MIND)

The MIND descriptor is formally defined as :

$$\mathrm{MIND}(I, \mathbf{x}_c, \mathbf{r}) = \frac{1}{n}\exp(-\frac{D_p(I, \mathbf{x}_c, \mathbf{x}_{c+r})}{V(I, \mathbf{x}_c)}) \quad \mathbf{r}\epsilon R \tag{1}$$

in which D_p in (3) is a similarity measure between small patches around a center voxel $\mathbf{x}_c$ respectively a neighbouring voxel $\mathbf{x}_{c+r}$, both of which are taken from a neighbourhood R. Effectively, MIND yields a vector of size R which represents the local structure information around each center voxel. In (1) n and $V(I, \mathbf{x}_c)$ are normalization terms with n the neighbourhood size and:

$$V(I, \mathbf{x}_c) = \frac{1}{6}\sum_{r\epsilon N} D_p(I, \mathbf{x}_c, \mathbf{x}_{c+r}) \tag{2}$$

Furthermore, the similarity measure D_p is given by:

$$D_p(I, \mathbf{x}_c, \mathbf{x}_{c+r}) = \sum_{p\epsilon P}(I(\mathbf{x}_c + p) - I(\mathbf{x}_{c+r} + p))^2 \tag{3}$$

representing the sum of squared difference (SSD) between two identically shaped patches P.

In essence, MIND maps each voxel onto a self-similarity vector that embeds the local structure. Subsequently, the registration problem is solved by minimizing the sum of absoluted differences of MIND measures over two images I and J under a certain deformation function.

2.2 Definition of the Hessian and Reorientation Strategy

An alternative method to measure local structure is based on the Hessian (HE):

$$HE = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \tag{4}$$

The calculation of image derivatives is based Gaussian kernals (width: σ_g). Typically, the HE is sensitive to the linelike structures (even structures). It captures local orientaion which is also covered by MIND.

In Diffusion Tensor Image (DTI) registration, tensor reorientation has been incorporated in the registration optimization [14]. Thereby the local orientation can be preserved during the registration and registration efficiency also might be improved. A standard way to do so is the Finite strain (FS) strategy in which a rigid rotation component is calculated by decomposing the deformation gradient:

$$R(\mathbf{x}) = (J(\mathbf{x})J(\mathbf{x})^T)^{-\frac{1}{2}} J(\mathbf{x}) \tag{5}$$

where $R(x)$ in (5) is the rotation matrix in a voxel and $J(x)$ is deformation gradient (i.e. Jacobian) at that voxel. Subsequently, the diffusion tensor $T(x)$ is reoriented by:

$$T'(\mathbf{x}) = R^T(\mathbf{x})T(\mathbf{x})R(\mathbf{x}) \tag{6}$$

Our structural representations make that a reorientation strategy is directly applicable in a new optimization framework. The effect of reorientation on a HE image is illustrated in 2D in Fig.1. It can be seen in the HE image without reorientation (b) that there is a disparity with the 'ground truth' (e), see the white arrows. Instead, the reoriented image (c) from (b) much closer resembles the ground truth (f).

2.3 Registration Framework

Inspired by [10] our similarity metric at each voxel is the mean absolute differences of HE's:

$$\mathcal{S}(HE_f(\mathbf{x}), HE_m(\mathbf{x})) = \text{mean}|HE_f(\mathbf{x}) - HE_m(\mathbf{x})|^2 \tag{7}$$

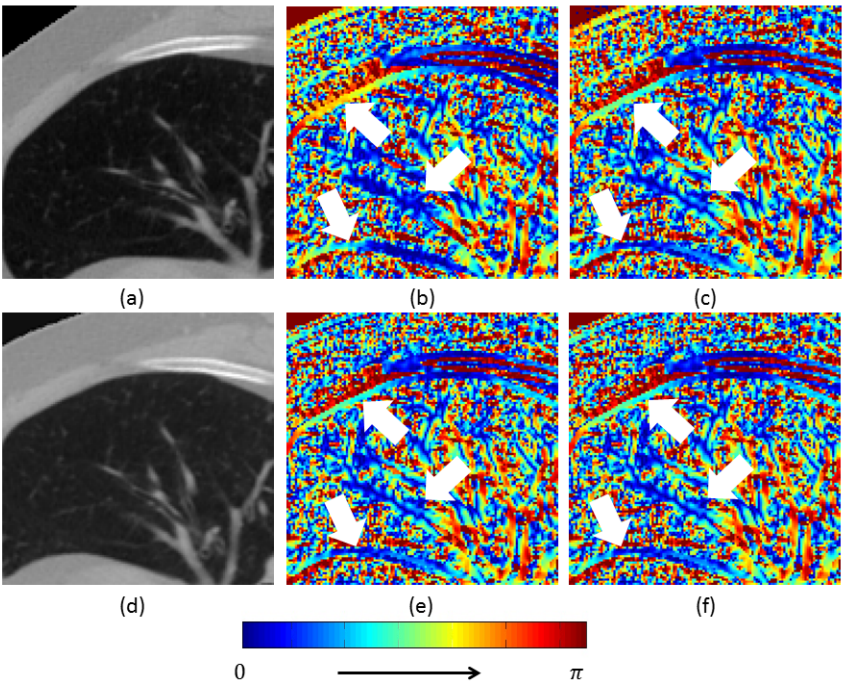


Fig. 1. Illustration of the effect of reorientation on the HE. (a) is an example lung CT image that we rotated by 30 degrees clockwise around the z direction yielding image (d). Images (b)(c)(e)(f) display the orientation of the first eigenvector derived from the Hessian; (b) shows a 30 degrees rotated tensor image derived from image (a) ($\sigma_g = 0.5$ calculated without reorientation) and (e) is the 'ground truth' calculated from (d). (c) is the result after rotation *and* reorientation from (a) and (f) is a duplicate version of (e) just for comparison.

where the $HE_f(\mathbf{x})$ and $HE_m(\mathbf{x})$ are the Hessian in voxel $\mathbf{x}$ from the fixed image and moving image respectively. We map the fixed image and moving image to Hessian space and a Gaussian Newton optimisation scheme was used to minimize following the cost function:

$$\underset{\mathbf{u}}{\text{argmin}} = \sum_{\mathbf{x}} \mathcal{S}(HE_f(\mathbf{x}), R(\mathbf{x}+\mathbf{u})^T HE_m(\mathbf{x}+\mathbf{u}) R(\mathbf{x}+\mathbf{u}))^2 + \alpha \text{tr}(\nabla \mathbf{u}(\mathbf{x})^T \nabla \mathbf{u}(\mathbf{x}))^2 \tag{8}$$

where $R(\mathbf{x}+\mathbf{u})$ is the rotation matrix , $\mathbf{u}$ is the deformation field fand α is a parameter that weights a regularisation term.

In the optimizatin step, MIND is recalculated after a certain number of deformation steps (more detail can be found in [10]). In contrast, our approach involving the Hessian enables to incorporate a reorientation term into the cost function and thus cope with the deformation (Equation (8)). The Hessian is only calculated once and the deformation of Hessian is embedded in the optimization.

Table 1. Mean distances(mm) of landmarks in 10 lung CT images: INITIAL is prior to registration; HE give the outcome based on HE registration without reorientation; ROHE are the result if reorientation is included; MIND is the outcome using the MIND framework

Case	1	2	3	4	5	6	7	8	9	10	Mean
INITIAL	4.07	4.40	7.03	9.91	7.51	10.99	11.13	15.06	8.02	7.43	8.56
HE	1.91	1.80	2.30	2.60	3.02	5.13	4.74	9.53	3.14	2.67	3.68
MIND	1.05	1.06	1.23	1.48	1.62	1.61	2.04	3.46	1.37	1.63	1.66
ROHE	1.08	1.06	1.27	1.53	1.56	1.64	1.92	3.26	1.36	1.60	1.62

As such, we can already reckon with an altering self-similarity when determining Gauss-Newton optimization steps (given by the gradient of Equation (8)). By doing so it may be expected that the technique is less prone to convergence in local minima. The interpolation inherent to the registration problem was performed trilinearly in the derivative spaces after which an interpolated Hessian was calculated.

3 Experiments and Results

3.1 DIR-Lab 4D CT

We tested our registration framework based on the HE on the public dataset provided by the DIR-Lab at the University of Texas [15]. This data set consists of thorax CT volumes acquired in inspiration as well as expiration from 10 subjects in which 300 landmarks were annotated by experts. A comparison of our approaches with the MIND approach are collated in Table 1. Thereby, we used $\sigma_g = 0.5$ and $\alpha = 0.1$, all of which were chosen to be comparable to the MIND approach. .

A two sample t-test was used to compare the registration strategies. Particularly, the results of ROHE and MIND are significant improvements over the initial situation as well as registration based on HE without reorientation (all: $p<0.05$) . ROHE has lower mean distance than MIND, but this difference is not significant. Fig. 2 shows a registration case based on ROHE and MIND. Although the difference in performance is subtle, it can be clearly seen that near the lung boundary ROHE outperforms MIND (see red arrow). Here, MIND appears to converge in a local minimum since increasing the number of steps does not improve the registration outcome. We found similar results on other cases. Additionally, we calculated the sum of absolute intensity differences (SAD) for ROHE and MIND, see Table 2, which shows the same trend.

3.2 Pre-contrast and Post-contrast Abdominal MRI

We applied our algorithms to clinical datasets from [16]. From this dataset we sequentially included the first 5 patients diagnosed with Crohn's disease . Patients drank 1600ml of a hyperosmolar fluid (Mannitol, 2.5%, Baxter, Utrecht,

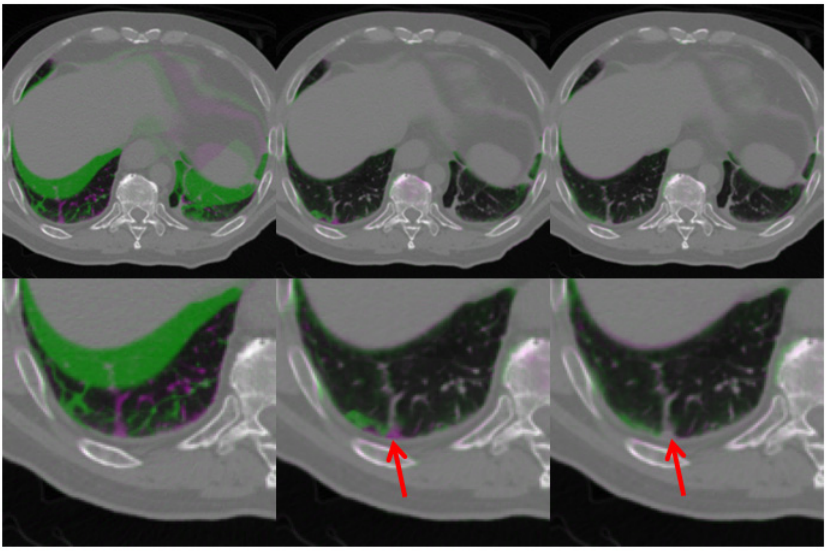


Fig. 2. Registration comparison for case 4. Top is an overview image, bottom shows a detail (as indicated). From left to right are images prior to registration, images registered by MIND and by ROHE, respectively. In all imageshe inhale phase (fixed image) is displayed in magenta and exhale phase (moving image) is displayed in green.

The Netherlands) 1 hour before acquiring the MRI scans to achieve bowel distention. MR imaging included a high resolution, 3D T1-weighted spoiled gradient echo sequence with fat saturation (pre-contrast MRI), followed by a free-breathing 3D+time Dynamic Contrast Enhanced (DCE)-MRI data acquisition on a 3.0T MRI scanner (Intera, Philips Healthcare, Best, The Netherlands) by a 3D spoiled gradient echo sequence. A contrast agent (Gadovist 1.0 mmol/ml, Bayer Schering Pharma, Berlin, Germany) was injected (0.1 ml/kg bodyweight) during the DCE-MRI acquisition. The DCE sequence was also succeeded by a high resolution, 3D T1-weighted spoiled gradient echo sequence with fat saturation (post-contrast MRI). A bowel relaxant (20 mg, Buscopan, Boehringer, Ingelheim, Germany) was administered to the subjects scans to minimize bowel movement. Registration of the pre and post contrast scans is considered as an important step to quantify disease severity by means of the relative contrast enhancement .

Table 2. Sum of absolute intensity differences ($\times 10^{10}$) prior to registration (INITIAL) and after registration based on MIND and ROHE, respectively

Case	1	2	3	4	5	6	7	8	9	10
INITIAL	6.82	11.03	12.86	10.15	10.75	25.29	29.44	44.46	20.83	23.84
MIND	0.84	1.35	1.15	1.54	1.34	5.92	5.76	**5.67**	3.08	4.15
ROHE	**0.78**	**1.25**	**1.07**	**1.36**	**1.23**	**5.47**	**5.10**	5.84	**2.76**	**3.71**

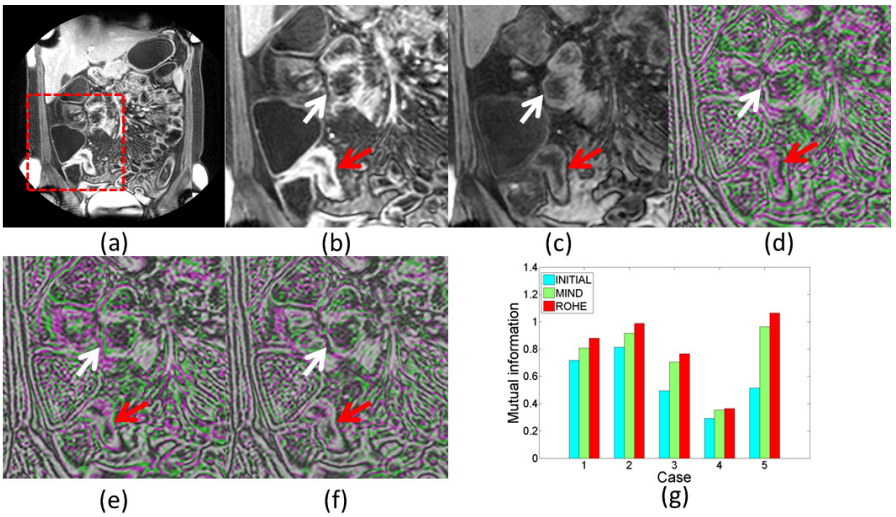


Fig. 3. Registration comparison on abdominal imaging data. (a) is a post-contrast MR image; (b) is sub-regions indicated in (a). (c) is the subregions from pre-contrast MR in same slice position as (b). (d) - (f) are the color maped versions of the sub-regions prior before registration (d), after registration by MIND (e), ROHE (f), respectively. The locally normalized post-contrast MR sub-region is displayed in magenta and the locally normalized pre-contrast sub-region is displayed in green. (g) quantifies the registration performance via the the mutual information.

The outcome of two registration approaches on a representative example are shown in Fig. 3. The terminal ileum containing Crohn's disease, is indicated by a red arrow. Fig. 3 (f) results after ROHE based registration yielding best outcome particularly around this region. What is more, Fig. 3 (g) quantifies the registration performance via the the mutual information on all cases (since manually indicated landmarks apeared irreproducible on this data). This figure also reflects that ROHE predominantly gives the best performance .

4 Conclusion

We presented a novel registration procedure based on the Hessian that incorporated a reorientation strategy into the registration optimization. The representation of the local self-smilarity as a tensor enabled keeping it up to date during deformation. As such, the registration procedure is efficient and not prone to fall in local minima. We showed that reorienting the hessian gave a significant improvement in registration accuracy ($p<0.05$) over leaving the reorientation out. Our technique also had a better performance over the existing MIND method on the DIR-Lab dataset as well as on abdominal MRI datasets albeit not significant. In the future we aim to differentiate in the weight given to varying structures (e.g. lines, planes, isotropic structures). Futhermore we will compare our method

with gradient based methods [11, 12]. Ultimately, we will use the technique to quantify Crohn's disease severity based on the relative contrast enhancement in registered images.

References

1. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality Image Registration by Maximization of Mutual Information. IEEE Transactions on Medical Imaging 16(2), 187–198 (1997)
2. Viola, P., Wells III, W.M.: Alignment by Maximization of Mutual Information. International Journal of Computer Vision 24(2), 137–154 (1997)
3. Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A.: Mutual-Information-Based Registration of Medical Images: A Survey. IEEE Transactions on Medical Imaging 22(8), 986–1004 (2003)
4. Loeckx, D., Slagmolen, P., Maes, F., Vandermeulen, D., Suetens, P.: Nonrigid Image Registration Using Conditional Mutual Information. IEEE Transactions on Medical Imaging 29(1), 19–29 (2010)
5. Gorbunova, V., Lo, P., Ashraf, H., Dirksen, A., Nielsen, M., de Bruijne, M.: Weight Preserving Image Registration for Monitoring Disease Progression in Lung CT. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 863–870. Springer, Heidelberg (2008)
6. Song, G., Tustison, N., Avants, B., Gee, J.C.: Lung CT Image Registration Using Diffeomorphic Transformation Models. In: van Ginneken, B., Murphy, K., Heimann, T., Pekar, V., Deng, X. (eds.) Medical Image Analysis for the Clinic: A Grand Challenge, pp. 23–32. CreateSpace, Charleston (2010)
7. Yi, Z., Soatto, S.: Multimodal Registration via Spatial-Context Mutual Information. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 424–435. Springer, Heidelberg (2011)
8. Zhuang, X., Arridge, S., Hawkes, D.J., Ourselin, S.: A Nonrigid Registration Framework Using Spatially Encoded Mutual Information and Free-Form Deformations. IEEE Transactions on Medical Imaging 30(10), 1819–1828 (2011)
9. Klein, S., van der Heide, U.A., Lips, I.M., van Vulpen, M., Staring, M., Pluim, J.P.: Automatic Segmentation of the Prostate in 3D MR Images by Atlas Matching Using Localized Mutual Information. Medical Physics 35(4), 1407–1417 (2008)
10. Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, S.M., Schnabel, J.A.: MIND: Modality Independent Neighbourhood Descriptor for Multi-Modal Deformable Registration. Medical Image Analysis 16(7), 1423–1435 (2012)
11. De Nigris, D., Collins, D., Arbel, T.: Multi-Modal Image Registration Based on Gradient Orientations of Minimal Uncertainty. IEEE Transactions on Medical Imaging 31(12), 2343–2354 (2012)
12. Li, B., Li, X., Wang, K., Qin, H.: Surface Mesh to Volumetric Spline Conversion with Generalized Polycubes. IEEE Transactions on Visualization and Computer Graphics 19(9), 1539–1551 (2013)

13. Heinrich, M.P., Jenkinson, M., Gleeson, F.V., Brady, S.M., Schnabel, J.A.: Deformable Multimodal Registration with Gradient Orientation Based on Structure Tensors. Annals of the British Machine Vision Association 5(2), 1–11 (2011)
14. Yeo, B.T., Vercauteren, T., Fillard, P., Peyrat, J.M., Pennec, X., Golland, P., Ayache, N., Clatz, O.: Dt-refind: Diffusion Tensor Registration with Exact Finite-Strain Differential. IEEE Transactions on Medical Imaging 28(12), 1914–1928 (2009)
15. Castillo, R., Castillo, E., Guerra, R., Johnson, V.E., McPhail, T., Garg, A.K., Guerrero, T.: A Framework for Evaluation of Deformable Image Registration Spatial Accuracy Using Large Landmark Point Sets. Physics in Medicine and Biology 54(7), 1849–1870 (2009)
16. Ziech, M., Lavini, C., Caan, M., Nio, C., Stokkers, P., Bipat, S., Ponsioen, C., Nederveen, A., Stoker, J.: Dynamic Contrast-Enhanced MRI in Patients with Luminal Crohn's Disease. European Journal of Radiology 81(11), 3019–3027 (2012)

Registration of Prone and Supine CT Colonography Datasets with Differing Endoluminal Distension

Holger R. Roth[1], Jamie R. McClelland[1], Thomas E. Hampshire[1], Darren J. Boone[2], Yipeng Hu[1], Marc Modat[1], Hui Zhang[1], Sebastien Ourselin[1], Steve Halligan[2], and David J. Hawkes[1]

[1] Centre for Medical Image Computing, University College London, WC1E 6BT, UK
h.roth@ucl.ac.uk
[2] Centre for Medical Imaging, Department of Radiology, University College Hospital, London, NW1 2BU, UK

Abstract. Robust registration between prone and supine data acquisitions for CT colonography is pivotal for medical interpretation but a challenging problem, especially in sub-optimally prepared patients. This paper introduces a prone and supine registration method that aims to overcome the difficulties posed by differences in luminal distension and bowel cleansing. The endoluminal surface is iteratively deformed using thin plate spline interpolation in order to increase similarity between prone and supine surfaces. Iterative deformation allows the re-computation of surface curvatures and, therefore, surface features to resemble one another more closely. Therefore, the similarity between surfaces increases with each optimization step when running a subsequent intensity-based registration in cylindrical space. Improved spatial alignment of endoluminal surfaces and better registration accuracies are shown in a limited number of challenging cases.

Keywords: Abdominal imaging, CT colonography oncology applications, registration, computed tomography, computer-aided diagnosis.

1 Introduction

CT colonography (CTC) 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 routine practice to obtain CT data with the patient both prone and supine to redistribute colonic residue and gas; fecal residue tends to move, whereas fixed mural pathology does not. Thus, matching corresponding locations between prone and supine acquisitions is pivotal for accurate interpretation [1]. Unfortunately, the colon undergoes three-dimensional deformations during patient repositioning [2] which complicates the interpretative task and can induce reader error.

Many algorithms have been reported that attempt to facilitate matching of corresponding locations between prone and supine acquisitions: Matching of distances along colonic centerlines enables navigation to approximate locations in

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 29–38, 2013.

both datasets [12,1,7,15,8,13]. Alternative methods that aim to register the full endoluminal surfaces have been proposed [10,11,4,16]. However, studies routinely use optimally cleansed and distended CTC datasets; yet in reality, approximately 50% of cases are "poorly prepared" [3]. Poor preparation may be due to incomplete cleansing or inadequate distension but a further, less quantifiable problem encountered in daily practice is that of markedly discrepant distension, ofen due to the redistributiuon of gas during patient repositioning. Registration algorithms should be able to handle sub-optimally prepared CTC datasets including those which demonstrate considerably different luminal distension between the prone and supine colonographic scans. In this paper we propose a novel method that aims to overcome limitations of CTC cases with marked differences in distension.

2 Methods

2.1 Prone and Supine Registration

We build upon methods that have been described previously [9] that establish full spatial correspondence between the prone and supine endoluminal surfaces. The entire surface is mapped to a cylinder utilizing a conformal mapping method based on Ricci flow [5]. The original surface curvature information is preserved during this step. Initialization is provided by robust haustral folds matching between both views [2]. Full surface correspondence is then achieved using a non-rigid, cylindrical version of the B-spline registration. Registration is driven by local shape measurements, i.e. shape index (SI) computed on the colon surface. The sum-of-squared differences (SSD) of these SI measures are used to drive the cylindrical registration [9]. After convergence of the algorithm, any point on the 3D surface can be mapped between both CTC acquisitions. Figure 1 illustrates the principle of this registration method.

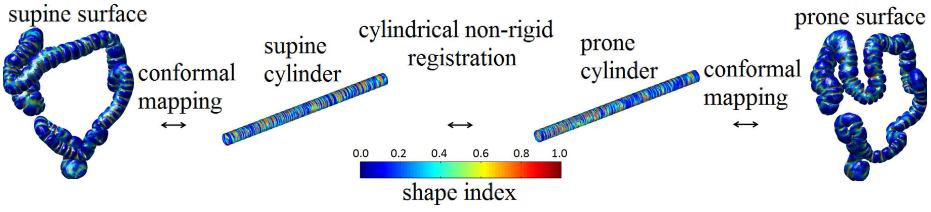


Fig. 1. Prone and supine registration of colon surfaces in cylindrical space. The color coding indicates the local shape index (SI) measurements [9].

The method as described in [9,2] performs an alignment of SI images in cylindrical space. The SI in both prone and supine is computed only on the original 3D colon surfaces, while registration aims to align SI features spatially in cylindrical space. The method performs well if the colon is well-cleansed and

similarly distended in both acquisitions. However, if the colon varies in distension between the two views or segmentation artifacts are present, a purely spatial alignment of features in the cylindrical images is more challenging. As features appear markedly different, registration results in an inferior alignment based on the intensity differences in the cylindrical images.

In order to overcome this limitation, we propose a new method that deforms one colon surface in an iterative manner in 3D in order to increase similarity between both surfaces. This allows re-computation of surface curvatures and therefore the SI values on the deformed surface at each iteration. The deformed surface aims to represent the target surface more closely. Therefore, the similarity of the cylindrical representations increases with each optimization step. This can subsequently improve the spatial alignment of surface features using a non-rigid image registration in cylindrical space. Figure 2 illustrates this principle.

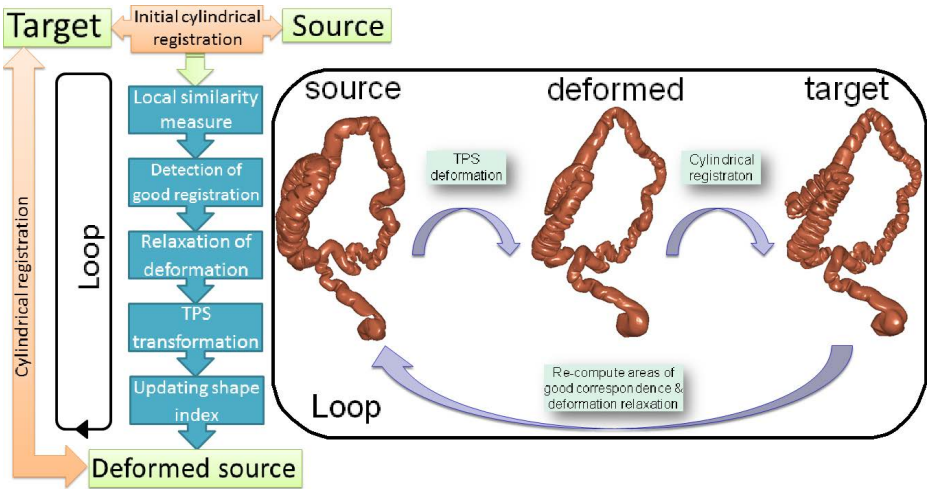


Fig. 2. Principle of the proposed iterative thin plate spline deformation process

2.2 Finding Well-corresponding Surface Areas

At each iteration step, areas of good local correspondence are determined based on the current registration result. The quality of local correspondence is measured using a local similarity measure S computed with a certain window-sized area of the endoluminal surface (as illustrated in Figure 3). We used a multi-resolution approach with local windows of different sizes and the negative SSD as the local similarity measure. A threshold t_s then separates good quality areas of the registration from inferior quality regions based on the window sizes and resolutions. A suitable value for t_s was found empirically and held constant for all cases in this study ($t_s = -0.025$).

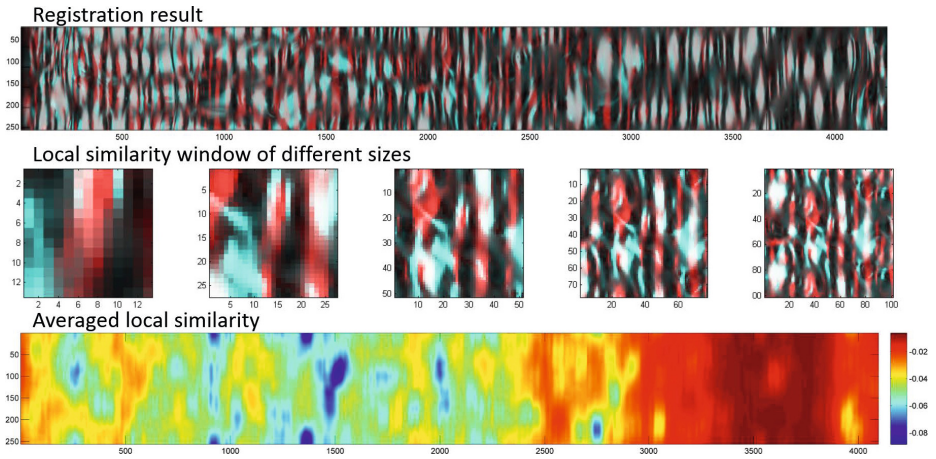


Fig. 3. Finding well-corresponding surface: The deformed source and target cylindrical images are shown as an overlay in red and cyan respectively (top row). Red and cyan pixels of the same intensity are displayed in grey. The registration result is used to compute local similarity (SSD) windows of several resolution levels (middle row). The average of these local windows is used to determine areas of good registration quality (bottom row).

2.3 Deformation Relaxation Approach

We ultimately want to re-compute the cylindrical non-rigid registration in order to account for the rotational twist and individual fold deformations in cylindrical space. However, in order to correct for misaligned areas in cylindrical space caused by dissimilar features from previous registration, the B-spline deformation grid is 'relaxed' in areas where the registration was poorly aligned, judging by the previous local similarity measure ($S < t_s$). This 'relaxation' of the cylindrical deformation field is achieved by utilizing the bending energy term in the optimization function of the cylindrical B-spline registration formulation [9]: All bad areas ($S < t_s$) of the registration are ignored when computing the similarity measure in this formulation. Therefore, in these areas, the registration algorithm is solely minimizing the bending energy of the deformation field. Whereas, features in good areas ($S \geq t_s$) are aligned using the SSD measure in the optimization of the cylindrical B-spline registration. This modified registration provides a 'relaxed deformation field' which can be used to re-initialize the cylindrical registration scheme over the full endoluminal surface for the next iteration step. This deformation relaxation is computed at the final resolution level of an image pyramid as the initial cylindrical registration has already accounted for large deformations. We are really only interested in capturing the features in the vicinity of good regions ($S \geq t_S$) in order to improve the

overall registration result. Figure 4 illustrates this relaxation of the deformation field in the bad areas of the cylindrical images. The result of this cylindrical registration is then used to initialize the next iteration of the proposed algorithm as illustrated in the flow chart of Fig. 2.

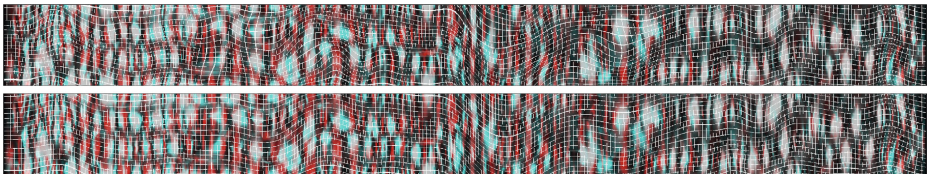


Fig. 4. Top: Deformation field from initial registration result. Bottom: Deformation after relaxation. This step keeps the well-corresponding areas of the deformation but "relaxes" the control points elsewhere. The deformation field has been used to deform a regular grid and been overlaid on top of the deformed source (red) and target (cyan) cylindrical images in both cases. Red and cyan pixels of the same intensity are displayed in grey after registration.

2.4 Surface Deformation Using Thin Plate Splines

In order to deform one surface, a thin plate spline (TPS) is fitted using a set of local point correspondences acquired. At each iteration step the majority of point correspondences are extracted from surface areas where the registration found good local correspondence ($S \geq t_S$). TPS allow the non-linear transformation of the whole image space of the source surface towards the space of the target surface. The transformation is based on physical principles. This enables the selection of different local bending properties of TPS in order to allow different degrees of local deformations. TPS can be fully fitted based on a set of point correspondences between the source and target 3D space and the bending properties as TPS can be estimated in a closed-form solution using a radial basis function $\varphi(r)$[14]. This function defines the transformation of any point x in the source space to $f(x)$ in the target space, given a set of correspondences functioning as control points of the TPS $\{w_i; i = 1, 2, ...K\} : f(x) = \sum_i^K c_i \varphi(\|x - w_i\|)$. For TPS the radial basis function is defined as $\varphi(r) = r^2 \log r$, while $\{c_i\}$ denotes a set of transformation coefficients. Additionally, an energy term is minimized in order to allow smooth TPS fitting:

$$E = \iiint \left[\left(\frac{\partial^2 f}{\partial x}\right)^2 + \left(\frac{\partial^2 f}{\partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial z}\right)^2 + 2 \times \left[\left(\frac{\partial^2 f}{\partial xy}\right)^2 + \left(\frac{\partial^2 f}{\partial yz}\right)^2 + \left(\frac{\partial^2 f}{\partial xz}\right)^2 \right] \right] dx\ dy\ dz \tag{1}$$

Surface areas with $S \geq t_s$ are then used to extract a dense set of local correspondences, allowing for increased bending of the TPS transformation. In contrast,

surface areas with $S < t_s$ are used to extract a sparser set of point correspondences, allowing for less bending. This is specified by specifically weighing the influence of each corresponding point using different scalar factors. Figure 5 shows the extraction of corresponding point locations between the source and target colonic endoluminal surface which are used to fit the TPS at each iteration step. The right hand side shows better local correspondences and therefore a much higher concentration than in the middle section.

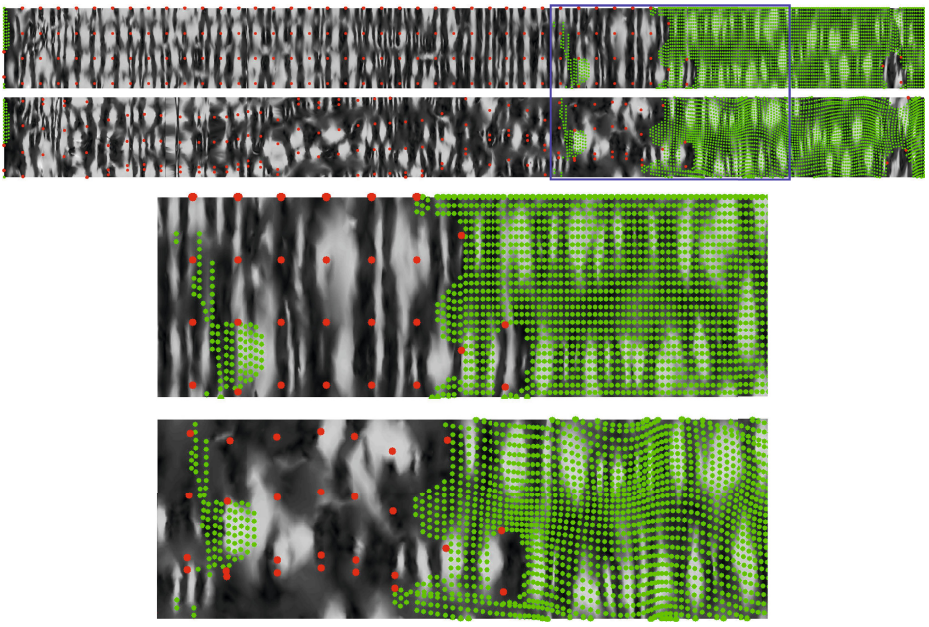


Fig. 5. Extraction of points of good (densely sampled green dots) and bad correspondences (sparser sampled magenta dots). The cylindrical images show the shape index intensities of the source (top) and target (bottom) colon surfaces. The marked region is enlarged for better visibility.

2.5 Re-establishing Correspondence on the Updated Surfaces

After applying the TPS deformation the source surface resembles the target surface more closely. In Fig. 6 a source surface is being deformed using TPS and mapped into the 3D space of the target image. Good overlap can be seen in the colonic segments where previously a good spatial correspondence was detected using the local similarity measure S (segment between arrows marked on the right colon in Fig. 6, right). The SI values can now be re-computed on the deformed surface and mapped onto the cylindrical source space using the known conformal mapping. This leads to more similarity between both cylindrical images. A subsequent registration in cylindrical space driven by these SI values is then more likely to align features correctly in the areas neighboring the good

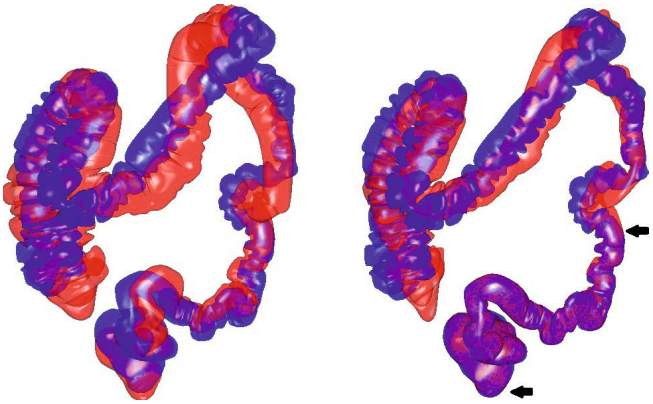


Fig. 6. Colon surface deformation using thin plate splines. Left: source (red) & target (blue) surfaces. Right: deformed source (red) & target (blue) surfaces. Good overlap can be seen in the colonic segments between the arrows where previously a good spatial correspondence was detected.

areas ($S \geq t_s$). This process is continued until the TPS is fitted using points over the full cylindrical image.

3 Results

The proposed algorithm was applied to four patient cases where the cylindrical non-rigid registration alone could not provide sufficient registration accuracy, e.g. average registration errors of over 10 mm. As part of this research, three experienced observers manually found corresponding haustral folds in the prone and supine scans using virtual fly-through renderings, the final reference standard was achieved in consensus. This reference standard was used to measure registration accuracy at corresponding haustral fold centre points.

Figure 7 shows the mean registration error for each case with increasing iterations of the proposed algorithm. Decreased registration error is observed in three cases (A, B and C) when previously misaligned features that neighbor good regions are 'captured' by the algorithm and then aligned correctly. The average registration error of all four cases improved from 14.5 mm to 13.2 (-9 %) mm; with the largest improvement in case C with 2.8 mm (-16.4 %) decrease in registration error. One case (D) did not improve but the initial error was already relatively small. For comparison, E shows an well-prepared case (which was not included in the evaluation) in which the registration error was low to start with due to the inital good similarity in distension [9]. Here, the method stops after fewer iterations and the error remains stable at 6.0 mm. This is because the similarity measure will already be high initially and cause the proposed method to terminate at an early iteration.

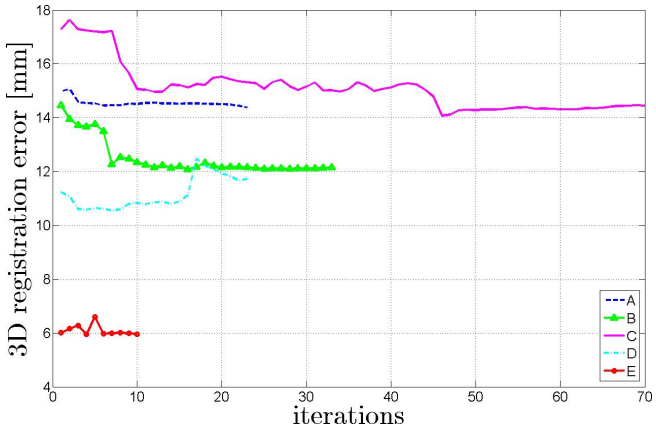


Fig. 7. Evolution of registration error in mm over several iterations of the proposed optimization method

4 Discussion

This paper introduced a prone and supine registration method that aims to overcome the difficulties posed by marked differences in distension between datasets. Differences in distension can cause endoluminal features to appear very different between prone and supine CT acquisitions. The method iteratively transforms one surface to resemble the other surface more closely, achieving improved registration accuracy in three out of four cases with large differences in distension.

Current stability issues can occur when re-sampling the transformed surface meshes in order to re-generate the cylindrical images. For example, case D shows an improvement in mean registration error for over 15 iterations in comparison to the initial value. However, the registration misaligned one substantial haustral fold in the cecum end of the image in the 16th iteration, causing an increasing step of mean registration error by 1.2 mm (11 %).

Local *SSD* might be a sub-optimal measurement to establish whether registration is successful or not. Correlation-based measures might be less affected by differences in the *SI* images. Potential improvement in estimation of surface curvatures could be achieved by using volumetric splines representing the colon surface [6]. An alternative to TPS when modeling the 3D transformations between prone and supine positions could be the use of B-splines. However, B-splines are constrained with control points that are defined over the entire three-dimensional space that embeds the colonic endoluminal surfaces. As the colon does not cover all that 3D space, an hierarchical multi-resolution approach or *Non-Uniform Rational B-splines* would be necessary. Alternatively, more realistic deformations could potentially be achieved with bio-mechanical models.

Acknowledgment. This work was undertaken at UCL and UCLH, which receives a proportion of funding from the *NIHR Biomedical Research Centre funding scheme* (Grant No. RP-PG-0407-10338) and the *EPSRC-CRUK Comprehensive Cancer Imaging Centre of UCL and KCL* (Grant No. C1519AO).

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), 332 (2001)
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. 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(5), 1076–1082 (2011)
4. 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(2), 551–560 (2007)
5. Jin, M., Kim, J., Luo, F., Gu, X.: Discrete surface Ricci flow. IEEE Trans. Vis. Comput. Graphics 14(5), 1030–1043 (2008)
6. Li, B., Li, X., Wang, K., Qin, H.: Surface mesh to volumetric spline conversion with generalized polycubes. IEEE Transactions on Visualization and Computer Graphics 19(9), 1539–1551 (2013)
7. 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(10), 2912–2923 (2004)
8. 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)
9. Roth, H.R., McClelland, J.R., Boone, D.J., Modat, M., Cardoso, M.J., Hampshire, T.E., Hu, M., Punwani, S., Ourselin, S., Slabaugh, G.G., Halligan, S., Hawkes, D.J.: 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`
10. Suh, J.W., Wyatt, C.L.: Registration of prone and supine colons in the presence of topological changes. In: SPIE Med. Imaging: Physiology, Function, and Structure from Medical Images, vol. 6916(1), p. 69160C (2008)
11. Suh, J., Wyatt, C.: Deformable registration of supine and prone colons for computed tomographic colonography. J. Comput. Assist. Tomogr. 33(6), 902–911 (2009)
12. 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(5), 1296–1304 (2009)
13. 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(945), 740–744 (2006)

14. Wahba, G.: Spline models for observational data, vol. 59. Society for Industrial Mathematics (1990)
15. 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(12), 5595–5603 (2009)
16. Zeng, W., Marino, J., Chaitanya Gurijala, K., Gu, X., Kaufman, A.: Supine and prone colon registration using quasi-conformal mapping. IEEE Trans. Vis. Comput. Graphics 16(6), 1348–1357 (2010)

Spatial Correspondence between Prone and Supine CT Colonography Images: Creating a Reference Standard

Thomas Hampshire[1], Emma Helbren[2], Andrew Plumb[2], Holger Roth[1], Steve Halligan[1], and David Hawkes[1]

[1] Centre for Medical Image Computing, University College London, Gower Street, London WC1E 6BT, UK
thomas.hampshire.09@ucl.ac.uk
[2] Centre for Medical Imaging, University College London, 255 Euston Road, London WC1 2BU, UK

Abstract. Matching corresponding locations between prone and supine CT colonography (CTC) images is difficult due to colonic deformations that occur between patient repositioning. We propose a novel method to allow a set of interpreting readers to establish a reference standard by matching corresponding locations in the prone and supine acquisitions. Independent matching of haustral folds was carried out for 17 CTC datasets by three readers, with the final reference standard being achieved in consensus. This resulted in 1743 reference standard points which have been made publicly available, along with the original CTC data at http://cmic.cs.ucl.ac.uk/CTC.

Keywords: CT colonography, haustral fold, reference standard.

1 Introduction

Matching corresponding locations between prone and supine CT colonography (CTC) images is difficult due to colonic deformations that occur on patient repositioning. However, achieving spatial correspondence between these patient positions could be used to test the accuracy of registration algorithms, which attempt to achieve this correspondence automatically. We propose a novel method to allow a set of interpreting readers to achieve a reference standard in consensus. First, the correspondence problem is reduced to matching a discrete set of locations by extracting haustral folds using a graph cut method applied to a curvature-based metric applied to a surface mesh generated from segmentations of the colonic lumen. The resulting set of locations can then be matched, one-to-one, between the prone and supine acquisitions. For this we have constructed a graphical user interface which provides visualisation of the interior endoluminal wall from three different aspects: external view, internal (virtual colonoscopy) views, and unfolded view. To the authors' knowledge, there has not been a published method for creating a reference standard in CTC, other than using the

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 39–45, 2013.

limited number of positions given by polyp location. Independent matching of haustral folds was carried out for 17 CTC datasets by three readers, with the final reference standard being achieved in consensus. This resulted in a total of 1743 reference points at corresponding fold pairs between the 17 prone and supine acquisitions. We have made this reference standard, along with the corresponding CTC data, available at `http://cmic.cs.ucl.ac.uk/CTC`, for use by the medical imaging community in algorithm development and testing.

2 Methods

2.1 Haustral Fold Segmentation

Haustral folds are first segmented using the method in [1]. Here we recognise that haustral folds are elongated ridge-like structures and can be identified by extracting the maximum k_1, and minimum k_2 curvature measurements from a triangular mesh representation of the colonic wall. Using the topology of the surface mesh, we treat it as a graph, with graph nodes defined by the mesh vertices and graph edges defined by the mesh edges. We then perform a graph cut segmentation to minimise the energy function:

$$\sum_p E_p(f_p) + \sum_p \sum_{q \in N_p \setminus q} \delta_{pq}(f_p, f_q), \tag{1}$$

where f_p and f_q are the binary labels of neighbouring nodes p and q, corresponding to fold or non-fold, and N_p is the neighbour set of p defined by the directly connected vertices. The function E_p is defined as:

$$E_p(f_p) = \begin{cases} k_1 - \gamma||k_2|| & \text{if } f_p \text{ is fold} \\ \gamma||k_2|| - k_1 & \text{if } f_p \text{ is non-fold} \end{cases}, \tag{2}$$

and δ_{pq} is a Potts energy function smoothing term. The centre of each fold is taken as the vertex with the shortest maximum distance to any vertex lying on the border of the segmented region. Visual results are displayed in Figure 1.

2.2 Haustral Fold Matching

Custom software was used to allow each reader to match the identified haustral folds between the prone and supine views. The Graphical User Interface (GUI) is displayed in Figure 2. The GUI is displayed twice on a dual-monitor system to allow visualisation of both prone and supine acquisitions simultaneously. The primary elements of the GUI are the *external view*, the *internal view* and the *unfolded view*. In each view, the positions of the detected haustral fold centres are displayed as white spheres. A fold may be labelled by selecting the sphere with the mouse in both the prone and supine views, using one of the GUI elements listed above. Using multiple views allows the reader to have a better confidence in finding corresponding locations between the two views. As a reader is confident in a correspondence between prone and supine haustral folds, a label is made which is displayed in each GUI element as a unique number. Each GUI element is described in more detail in the following sections.

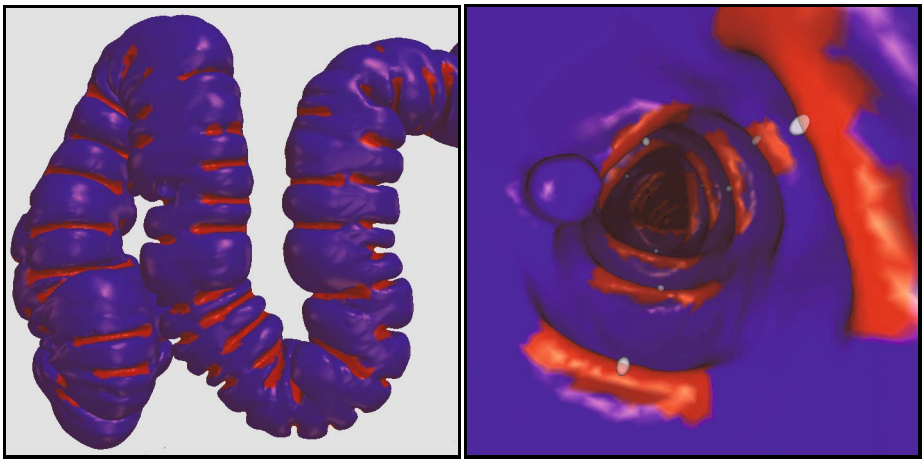


Fig. 1. External (right) and internal (left) views of automatically segmented haustral folds with marked centres. Red and blue sections represent fold and non-fold labelled vertices respectively.

External View. The external view displays the CTC surface rendering from a viewpoint positioned on the outside of the colon. This view proves useful in giving the readers an additional means of gaining spatial awareness, aiding in matching neighbouring haustral folds. Standard navigation controls allow for interactive manipulation of the camera: rotation, zooming, panning and spinning. A feature of this GUI element is the rendering of the CTC surface mesh using *frontface culling*. Face culling is usually utilised in the graphical rendering pipeline in the form of backface culling. This step determines if a polygon normal is facing away from the camera, in which case it is not displayed which makes the rendering process more efficient. This method can be used as there is generally no need to render the polygons on the sides of the object facing away from the camera, as it is occluded by the sides facing the camera. Frontface culling performs the reverse operation, by not rendering the polygons that are facing towards the camera. This is achieved by discarding all polygons whose dot product of their surface normal and the camera-to-polygon vector is less than or equal to zero. When viewing CTC surfaces, this allows the visualisation of the interior of the colon without occlusion from the exterior surface.

Internal View. The internal view displays the CTC surface rendering from a position in the interior of the colon lumen. This visualisation method is commonly known as 'virtual colonoscopy'. Navigation controls allow for camera translation along the colon lumen centreline, as well as rotation and spinning.

Unfolded View. The unfolded view displays the entire CTC surface rendering mapped onto a 2D rectangular plane using the conformal mapping method described in [2]. The image is display with the caecum on the left, and the

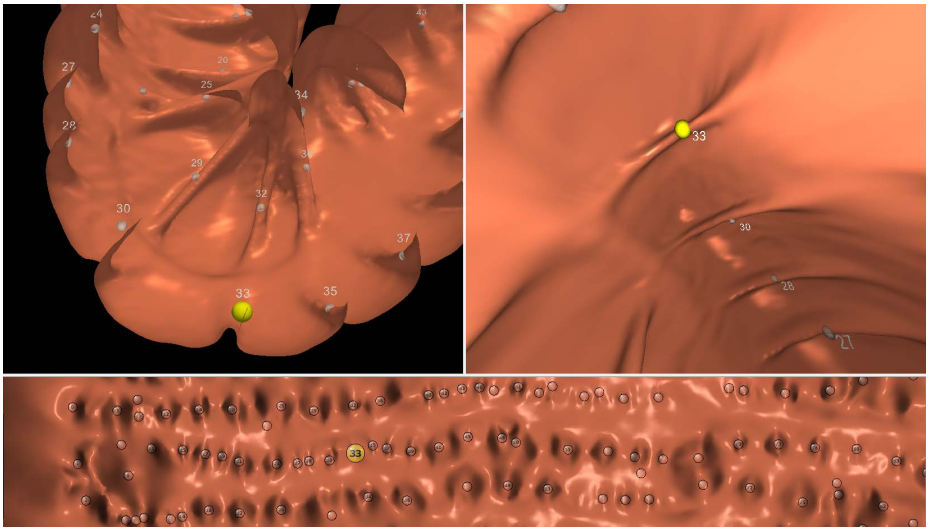


Fig. 2. Graphical User Interface (GUI) used for haustral fold matching. Top left shows the *external view*, top right shows the *internal view* and bottom shows the *unfolded view*. The white spheres on each view show the identified haustral fold positions. Numerical labels give a reference to a corresponding matching in the other patient position.

rectum on the right. The rectangular plane is a representation of an object with cylindrical topology, and therefore the top and bottom of the image correspond to neighbouring positions on the original surface mesh. The image is rendered using a phong shading model [3], with an augmented diffuse component. For each pixel, the viewing point and light can be assumed to be at the corresponding centreline position. This is calculated by finding the vertical line of pixels at the same horizontal position as the pixel to render, and taking the mean of the corresponding 3D positions on the original surface mesh. Each pixel intensity can then be calculated as:

$$I_p = k_a i_a + \frac{k_d(\hat{C} \cdot \hat{N} + 1) i_m}{2}^2 + k_s(\hat{R} \cdot \hat{C})^\alpha i_m \tag{3}$$

where k_a, k_d, k_s are the ambient, diffuse and specular reflection constants for the material; α is the material shininess constant; i_a, i_d, i_s are the RGB values of the incoming light source; $\hat{C}$ is the direction vector from surface point to centreline position; $\hat{N}$ is the surface normal at this point; and $\hat{R}$ is the direction that a reflected ray of light would take at this point, defined as:

$$\hat{R} = 2(\hat{C} \cdot \hat{N})\hat{N} - \hat{C} \tag{4}$$

The changes made to the diffuse component calculation allow for a polygon with normal facing away from the centreline position to still have an illumination value in the unfolded image.

2.3 Materials

Ethical approval and informed consent were obtained to use anonymised CT colonography data. Colonic cleansing and insufflations had been performed in accordance with current recommendations [4].

For the purpose of establishing a fold labelling between the prone and supine acquisitions, we selected the same 17 patient cases used in a previous publication [1,2]. In 12 of the cases, the colon was well distended in both views, and where fluid tagging (allowing for digital cleaning of residual fluid) was used or little fluid remained. This allowed a continuous segmentation over the full length of the colon using the methods described in [2]. The other 5 cases exhibit one or more area of local colonic collapse (see Table 1 for details).

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 from centreline end-point to centreline start-point. Locations of collapse are given (DC: descending colon; SC: sigmoid colon).

	Prone			Supine		
Case No.	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

A computer scientist (TH) with more than three years experience of reading CTC images, and two radiologists (EH, AP), independently established a reference standard by matching haustral folds using the GUI described above. Any folds where confident manual correspondences between the two views could not be established were excluded from the derived reference standard. All readers were unaware other readers' labelling results. After a minimal period of three weeks per case, the reference standards were compared for consistency, and any discrepancies were resolved by the three readers in consensus. This resulted in a total of 1743 corresponding fold pairs over the 17 validation datasets. This reference standard, along with the corresponding prone and supine CTC data sets, have been made publicly available at `http://cmic.cs.ucl.ac.uk/CTC`.

3 Results

The labelling of the three readers have been compared for consistency, with the results displayed in Table 2. Here the term *strong agreement* is used when, for a

fold in the prone data, all three readers have selected the same corresponding fold in the supine data. *Weak agreement* is used when two readers have selected the same supine fold for a given prone fold, and the third does not create a label. *Incomplete* is used when only a single reader labels a prone fold. *In conflict* indicates there is a conflicting label assignment for a given fold. Lastly, *empty* is used when no reader has labelled a particular fold. Labellings were marked as *resolved* if an agreement could be made in consensus, and *unresolved* if this was not true.

The final reference standard, achieved in consensus, had all unresolved fold matches removed, giving a total of 1743, or a mean of 102.5 per case. Comparisons between the individual reader labelling and the final reference standard are displayed in Table 3.

Table 2. Table displaying direct comparison of the three readers' reference standards

	Total	Resolved	Unresolved	% Resolved
Strong agreement	936	936	0	100.0%
Weak agreement	569	505	64	88.8%
Incomplete	530	212	318	40.0%
In conflict	130	87	43	66.9%
Empty	804	3	801	0.4%
Sum total	2969	1743	1226	58.7%

Table 3. Comparison of individual reader labelling against final reference standard

	Reader 1	Reader 2	Reader3
Total labels per reader	1666	1671	1449
Total labels agreeing with final RS	1491	1360	1296
Total labels disagreeing with final RS	3	56	34
Total labels in reader RS but not final RS	172	255	119
Total labels in final RS but not reader RS	249	327	413
Total labels in neither reader RS nor final RS	1054	971	1107

4 Conclusion

We present a novel method for achieving spatial correspondence between prone and supine CTC acquisitions by allowing a set of interpreting readers to create an accurate reference standard in consensus. The method was applied to 17 CTC patient cases, resulting in a reference standard of 1743 corresponding fold pairs. A large number of folds were not matched between acquisitions.

This is primarily due to the difficult task of matching corresponding folds. However, this method provides a mean number of 102.5 haustral fold matches per case which creates a much larger number of reference standard points per case for validating registration methods, than using polyps. The reference standard, along with the original CTC data has been made publicly available at `http://cmic.cs.ucl.ac.uk/CTC`.

References

1. Hampshire, T., Roth, R., Helbren, E., Plumb, A., Boone, D., Slabaugh, G., Halligan, S., Hawkes, D.: Endoluminal surface registration for CT colonography using haustral fold matching. Medical Image Analysis (2013)
2. 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`
3. Phong, B.: Illumination for computer generated pictures. Communications of the ACM 18(6), 311–317 (1975)
4. Taylor, S., 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)

Registration of Temporally Separated CT Colonography Cases

Holger R. Roth[1], Emma Helbren[2], Thomas E. Hampshire[1], Perry J. Pickhardt[3], Steve Halligan[2], and David J. Hawkes[1]

[1] Centre for Medical Image Computing, University College London, WC1E 6BT, UK
h.roth@ucl.ac.uk
[2] Centre for Medical Imaging, Department of Radiology, University College Hospital, London, NW1 2BU, UK
[3] Department of Radiology, University of Wisconsin Medical School, Madison, WI 53792-3252, USA

Abstract. Robust registration between prone and supine data acquisitions for CT colonography (CTC) is a useful tool for assessing clinically significant changes but a challenging problem. This is especially the case for polyp follow-up when scans are temporally separated. We investigated the ability of automatic registration to align CTC cases, acquired several months apart. 26 initial and follow-up cases were investigated and registration measured using the locations of 35 polyps in all available scans. Robust non-rigid feature-based initialization allowed registration of prone and supine CTC scans from patient cases not only acquired on the same day but also when acquired several months apart. A mean registration error of 17.4 (std. dev. 12.1) mm (median 14.9 mm, range 1.7 to 49.7 mm) was achieved when transforming polyp locations between longitudinal scans. The level of accuracy achieved was similar to previous studies that aligned CTC images acquired at the same sitting. Automatic registration of follow-up CTC investigations could be a useful adjunct for radiologists interpreting CTC for surveillance of colonic polyps.

Keywords: Abdominal imaging, CT colonography, follow-up investigations, registration, oncology applications, computed tomography, computer-aided diagnosis.

1 Introduction

Follow-up CT colonography (CTC) scans are necessary when a polyp detected on initial CTC is relatively small and so left in-situ. This is done when the risk of resection during subsequent optical colonoscopy (OC) outweighs the risk of leaving the polyp in-situ and monitoring its growth. Polyp growth, if any, is monitored by sequential CTC, taken months or years later [6]. Clearly, it is essential that the radiologist can identify the polyp under surveillance in both the initial prone/supine and the follow-up prone/supine data sets. Manual matching

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 46–52, 2013.

of polyps across longitudinal CTCs can be even more challenging and time-consuming than manual matching between prone and supine scans taken on the same day.

This study investigates the ability of a recently reported registration method [8] to temporally align separated CTC cases, acquired several months apart. No other study has investigated registration accuracy for methods that establish full surface correspondence between follow-up CTC examinations.

A polyp observed over several months is shown in Fig. 1 in coronal CTC views and in Fig. 2 using virtual fly-through renderings of the endoluminal surfaces.

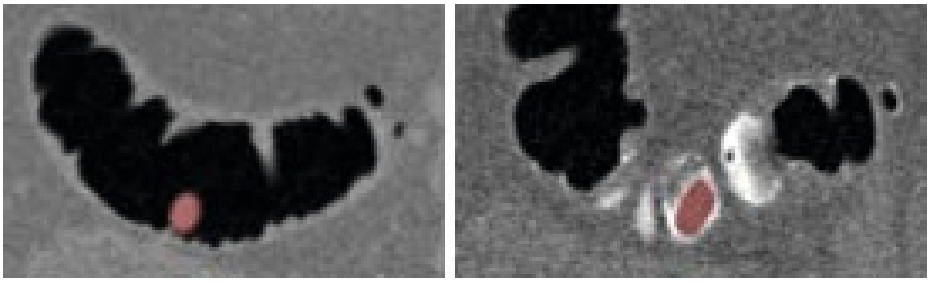


Fig. 1. Coronal views of a polyp in prone position (left) and prone position scanned 43 months later (right), highlighted using manual segmentation. The same polyp on the right is now covered by tagged fluids.

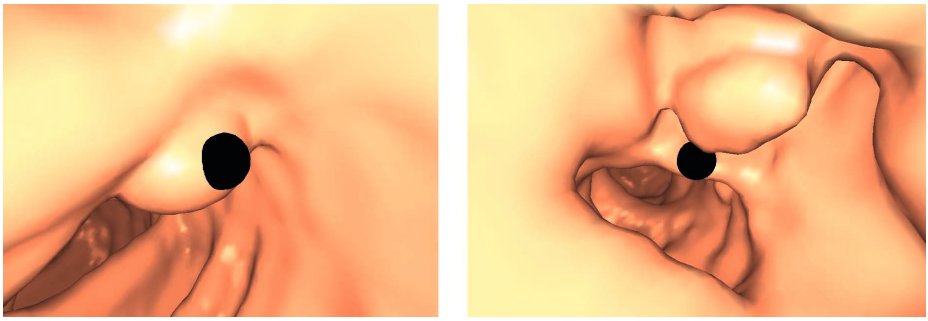


Fig. 2. Virtual fly-through renderings of a polyp in prone position (left) and prone position scanned 43 months later (right), after 'digital cleansing' of the tagged fluids. The polyps has now grown to about 11 mm in size. The black dot indicates corresponding locations using the registration result of the method described in [8].

2 Methods

2.1 Evaluation Data

Ethical approval and patient consent were obtained. All cases were selected from patients with two or more CTC investigations undertaken for the identification

and subsequent follow-up of colonic polyps. No attempt was made to select 'perfect cases' or exclude cases with poor distension from the study. These data had not been used previously for the development of the registration method. The evaluation sample consisted of 26 patients. From this group the radiologist (Emma Helbren) was able to identify 35 polyps present in both acquisitions in both the initial and subsequent CTC studies.

2.2 CT Colonography Registration

We build upon methods described previously [8,3] that establish full spatial correspondence between prone and supine endoluminal surfaces. The entire colon surfaces extracted from the initial and follow-up prone/supine CTC scans are mapped to cylinders utilizing a conformal mapping method based on Ricci flow [4]. The original surfaces' curvature information is preserved during this step. Initialization is provided by robust haustral fold matching between all four cylindrical views [3,2]. Full surface correspondence is then achieved using a non-rigid, cylindrical version of a B-spline registration method [5]. Registration is driven by local shape measurements, i.e. shape index (SI) computed on the colon surface. The sum-of-squared differences (SSD) of these SI measures are used to drive the cylindrical registration [8]. After convergence of the algorithm, any point on the 3D surface can be mapped between both CTC acquisitions. Figure 3 illustrates the principle of this registration method. Fig. 3 further illustrates how correspondence between all data sets of a follow-up study can be achieved. The follow-up prone (P) and supine (S) data sets acquired on the first or second occasion are superscripted with $_1$ or $_2$ respectively. The registration allows the transformation of any surface location between all temporally separated (longitudinal) and same-day acquisitions. One could transform points between all data sets by only computing three registrations: $P_1 \rightarrow P_2$, $S_1 \rightarrow P_1$ and $P_2 \rightarrow S_2$. Therefore,

$$S_1 \rightarrow S_2 = S_1 \rightarrow P_1 \circ P_1 \rightarrow P_2 \circ P_2 \rightarrow S_2. \quad (1)$$

However, $S_1 \rightarrow S_2$ is also computed in order to reduce any accumulated error that would occur when composing three transformations as in equation 1. Furthermore, the computation of $S_1 \rightarrow S_2$ allows the generation of a "consistency error" over the whole colonic surface. This would measure how similar the registration results are (e.g. resulting in the same anatomical correspondence) when transforming one point around the full transformation 'loop' ($S_1 \rightarrow S_2 \rightarrow P_2 \rightarrow P_1 \rightarrow S_1$). This might be a good indicator for judging the successfulness of the registration without referring to a reference standard at the polyps positions ('consistency registration error').

3 Results

Registration was performed on all 26 patients, and polyp locations in all subsequent acquisitions were estimated using the registration result. Table 1 lists the

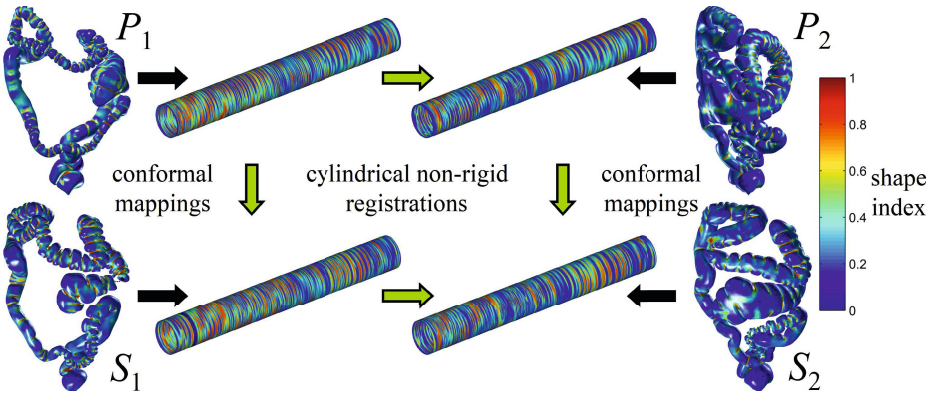


Fig. 3. Establishing correspondence between the follow-up prone (P) and supine (S) data sets acquired on the first or second occasion is superscripted with $_1$ or $_2$ respectively. Non-rigid registration of all colon surfaces is performed in cylindrical space after conformal mapping. The color coding indicates the local shape index (SI) measurements (see color scale) [8]. The different appearances of the 3D endoluminal surfaces between each scan illustrate the challenge of this registration task. A short section of endoluminal collapse (dotted line) is visible in the ascending colon of S_1. This translates to a 'ring' of missing data in the cylindrical representation of S_1 (left hand side of cylinder).

number of days separating each CTC study together with *longitudinal* and *consistency registration errors*. Using direct *longitudinal* transformations over time ($P_1 \rightarrow P_2$ and $S_1 \rightarrow S_2$), a mean longitudinal registration error of 17.4 (std. dev. 12.1) mm (median 14.9 mm, range 1.7 to 49.7 mm) was achieved. All errors measured using the Euclidean distance between transformed polyp location and the location of the targeted polyp.

Measuring the 'consistency error' around the loop (predicting polyp location through all acquisitions from the specified location on the initial supine scan alone) a mean Euclidean registration error of 26.9 (std. dev. 20.8) mm (median 28.0, range 0.9 to 84.5 mm) was achieved. For comparison, the mean registration errors between prone and supine CTCs acquired on the same day was 16.9 (std. dev. 17.6) mm (median 13.8 mm, range 1.5 to 83.9 mm). There is no significant difference between *longitudinal* and *same-day registration errors* ($p = 0.451$[1]).

Both, the *longitudinal errors* and *consistency errors* are not correlated to the number of days between the two CTC studies with $p = 0.105$ and $p = 0.055$ respectively[2].

[1] Related-Samples Wilcoxon Signed Rank Test, 1% significance level.
[2] Two-tailed Pearson Correlation, 1% significance level.

Table 1. Registrations of follow-up studies on external CTC data. The number of days separating each colonography study are listed together with *longitudinal*, *consistency* and *same-day registration errors* (averaged over the number of polyps per case).

Patient #	Days #	Polyps*	Longitudinal Error [mm]	Consistency Error [mm]	Same-day Error [mm]
1	1500	3	23.6	42.7	21.6
2	734	1	12.9	56.4	83.9
3	741	1	34.5	44.1	21.5
4	779	1	1.7	2.9	1.5
5	730	1	5.2	7.7	1.9
6	1524	2	30.1	30.0	15.3
7	779	1	49.7	84.5	50.3
8	1905	1	17.5	28.8	8.9
9	742	1	18.1	8.7	20.1
10	755	1	5.4	10.1	1.8
11	1886	1	9.3	11.5	9.3
12	865	1	8.5	9.5	13.8
13	757	1	15.3	29.0	13.8
14	369	2	13.3	28.2	22.3
15	1842	2	4.3	3.3	8.4
16	1498	1	8.0	31.0	24.5
17	747	2	21.4	15.3	3.3
18	371	3	33.7	33.7	26.6
19	755	2	8.9	18.6	14.7
20	375	1	6.3	0.9	4.1
21	405	1	17.1	27.8	21.5
22	753	1	29.0	53.3	4.7
23	749	1	38.2	49.0	4.2
24	266	1	14.5	18.9	6.0
25	777	1	16.7	51.0	27.4
26	735	1	9.1	2.6	8.3
Mean	**898**	**1.3**	**17.4**	**26.9**	**16.9**
Std. dev.	**480**	**0.6**	**12.1**	**20.8**	**17.6**
Minimum	**266**	**1.0**	**1.7**	**0.9**	**1.5**
Maximum	**1905**	**3.0**	**49.7**	**84.5**	**83.9**

*The total number of polyps is 35.

4 Discussion

The challenge of automatically registering the endoluminal colonic surface acquired by CTC separated by several months or years is potentially more challenging than registration between scans taken during the same CTC sitting. It was previously demonstrated that the proposed registration algorithm can accurately match prone and supine datasets acquired on the same day [8]. In the present study we explored a wider application – the follow-up of polyps on subsequent CTC taken months and years later.

Temporal separation increases the chance of dissimilarity between bowel preparation, distension, and overall quality of CTC when comparing data sets, which might make automatic registration between these data sets more difficult. Despite this challenge, the level of accuracy achieved by the registration algorithm was similar to studies registering between prone and supine on the same day; and these results agree with previous studies aligning CTC images obtained at same-day investigations using data that reflects clinical practice, i.e. including collapsed regions [9,7,1]. For example, Boone et al. reported a polyp registration error (mean $\pm$ standard deviation) of 19.9 mm $\pm$ 20.4 mm in 51 CTC patients [1] and Suh and Wyatt reported an average registration error of 30.1 mm for four polyps in four CTC cases [9]. Registration errors of less than 100 mm could already be clinically useful when relating between CTC scans. For example, Summers et al. found an accuracy of 100 mm useful in linking CTC findings to optical colonoscopy (100 mm corresponds to one mark on a standard colonoscope) [10].

We achieved accuracies showing registration to be robust for lesion matching over time. Therefore, automatic registration could be a useful adjunct for those interpreting CTC for the follow-up surveillance of colonic polyps. Registration is likely to be especially helpful for follow-up of small polyps that are likely harder to locate and to identify without assistance. The fact that there is no dependency of *longitudinal errors* and *consistency errors* on the length of time between the initial and follow-up CTC studies highlights that registration errors are more likely caused by differences in distension and bowel-preparation than by any anatomical changes that might occur over the months and years between studies.

Further applications of follow-up registration could include automatic detection of structural abnormalities on the endoluminal surface. One could use the similarity measure of the registration cost-function not only to achieve alignment between colonic surfaces but also to automatically identify areas of dissimilarity that might be caused by abnormalities arising between follow-up scans. Furthermore, the deformation fields resulting from the registration could be used to estimate the growth or change of anatomical structures such as polyps.

Acknowledgment. This work was undertaken at UCL and UCLH, which receives a proportion of funding from the *NIHR Biomedical Research Centre funding scheme* (Grant No. RP-PG-0407-10338) and the *EPSRC-CRUK Comprehensive Cancer Imaging Centre of UCL and KCL* (Grant No. C1519AO).

References

1. Boone, D.J., Halligan, S., Roth, H.R., Hampshire, T.E., Helbren, E., Slabaugh, G.G., McQuillan, J., McClelland, J.R., Hu, M., Punwani, S., Taylor, S.A., Hawkes, D.J.: CT colonography: External clinical validation of an algorithm for computer-assisted prone and supine registration (2013)
2. Hampshire, T.E., Roth, H.R., Boone, D.J., Slabaugh, G., Halligan, S., Hawkes, D.J.: Prone to supine CT colonography registration using a landmark and intensity composite method. In: Yoshida, H., Hawkes, D., Vannier, M.W. (eds.) Abdominal Imaging 2012. LNCS, vol. 7601, pp. 1–9. Springer, Heidelberg (2012)

3. Hampshire, T., Roth, H.R., Helbren, E., Plumb, A., Boone, D., Slabaugh, G., Halligan, S., Hawkes, D.J.: Endoluminal surface registration for CT colonography using haustral fold matching. Medical Image Analysis (2013)
4. Jin, M., Kim, J., Luo, F., Gu, X.: Discrete surface Ricci flow. IEEE Trans. Vis. Comput. Graphics 14(5), 1030–1043 (2008)
5. 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(3), 278–284 (2010)
6. Pickhardt, P., Hassan, C., Laghi, A., Zullo, A., Kim, D., Iafrate, F., Morini, S.: Small and diminutive polyps detected at screening CT colonography: a decision analysis for referral to colonoscopy. American Journal of Roentgenology 190(1), 136–144 (2008)
7. Roth, H.R., Boone, D.J., Halligan, S., Hampshire, T.E., McClelland, J.R., Hu, M., Punwani, S., Taylor, S., Hawkes, D.J.: External clinical validation of prone and supine CT colonography registration. In: Yoshida, H., Hawkes, D., Vannier, M.W. (eds.) Abdominal Imaging 2012. LNCS, vol. 7601, pp. 10–19. Springer, Heidelberg (2012)
8. Roth, H.R., McClelland, J.R., Boone, D.J., Modat, M., Cardoso, M.J., Hampshire, T.E., Hu, M., Punwani, S., Ourselin, S., Slabaugh, G.G., Halligan, S., Hawkes, D.J.: 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
9. Suh, J.W., Wyatt, C.L.: Registration of prone and supine colons in the presence of topological changes. In: SPIE Med. Imaging: Physiology, Function, and Structure from Medical Images, vol. 6916(1), p. 69160C (2008)
10. 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(5), 1296–1304 (2009)

A Classification-Enhanced Vote Accumulation Scheme for Detecting Colonic Polyps

Nima Tajbakhsh[1], Suryakanth R. Gurudu[2], and Jianming Liang[1]

[1] Department of Biomedical Informatics, Arizona State University
13212 East Shea Boulevard, Scottsdale, AZ 85259
{Nima.Tajbakhsh,Jianming.Liang}@asu.edu
[2] Division of Gastroenterology and Hepatology, Mayo Clinic
13400 E. Shea Blvd., Scottsdale, AZ 85259
Gurudu.Suryakanth@mayo.edu

Abstract. Colorectal cancer most often begins as abnormal growth of the colon wall, commonly referred to as polyps. It has been shown that the timely removal of polyps with optical colonoscopy (OC) significantly reduces the incidence and mortality of colorectal cancer. However, a significant number of polyps are missed during OC in clinical practice—the pooled miss-rate for all polyps is 22% (95% CI, 19%–26%). Computer-aided detection may offer promises of reducing polyp miss-rate. This paper proposes a new automatic polyp detection method. Given a colonoscopy image, the main idea is to identify the edge pixels that lie on the boundary of polyps and then determine the location of a polyp from the identified edges. To do so, we first use the Canny edge detector to form a crude set of edge pixels, and then apply a set of boundary classifiers to remove a large portion of irrelevant edges. The polyp locations are then determined by a novel vote accumulation scheme that operates on the positively classified edge pixels. We evaluate our method on 300 images from a publicly available database and obtain results superior to the state-of-the-art performance.

Keywords: Optical colonoscopy, polyp detection, voting scheme, random forest, boundary classification.

1 Introduction

Colorectal cancer (CRC) is the second leading cancer killer in the United States with 50,830 estimated deaths in 2013 [1]. High level mortality of CRC mainly stems from its late diagnosis which comes with a poor five-year survival rate of 10% [2]. Optical colonoscopy is the current gold standard screening method for color cancer screening and prevention. The goal of colonoscopy is to detect and remove colonic polyps which are known as precursors to CRC. Figure 1 shows examples of colonic polyps. It has been shown that early detection of colonic polyps improves survival rate of CRC [3]. However, polyp detection with colonoscopy is not an infallible task and as evidenced by several clinical studies (e.g., [4–6]), a significant portion of flat and pedunculated polyps remain undetected during colon screening. High polyp detection rate, as pointed out in [7], requires a high level of attentiveness, alertness, and sensitivity to visual characteristics of polyps from colonoscopists and such qualities are only procured after years of practice and experience.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 53–62, 2013.

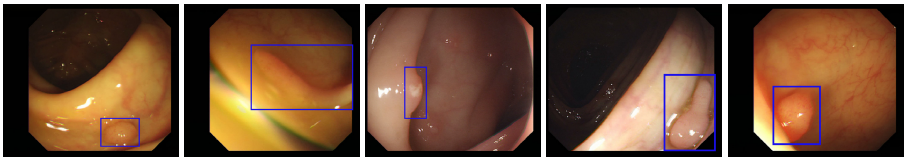


Fig. 1. Five colonoscopy images from the CVC database [8] with polyps highlighted by blue rectangles. As shown, polyps significantly vary in appearance.

Computer-aided polyp detection can be considered as a tool for reducing polyp miss-rate. The idea is to highlight regions with suspected polyps during colonoscopy procedures. Existing algorithms (e.g., [9–12]) for automatic polyp detection mainly use texture or shape information. The texture of a polyp becomes visible only if the camera captures close shots of the surface of the polyp. This condition is often met when polyps have already been detected by operators, which obviously eliminates the need for computer-aided detection. Shape information, on the other hand, appears to be more useful for polyp detection. However, raw shape information is affected by fragmented edge segmentation and in the absence of image context can mislead a detector towards irrelevant objects in the complex endoluminal scene.

Considering these observations, we propose a novel shape-based polyp detection method that aims to overcome the limitations of shape-based approaches. To reduce vulnerability against misleading objects, we propose a new image processing pipeline that filters out irrelevant boundaries in colonoscopy images by incorporating image context. The boundary removal mechanism mainly captures the changes in image appearance across polyp boundaries; it is therefore minimally affected by texture visibility limitations. To overcome the challenges posed by partial polyp segmentation, we propose a novel vote accumulation scheme that enables polyp localization from fragmented edge maps. We stress that our method is not designed to delineate polyps in images but to provide inexperienced colonoscopists with feedback on the locations of polyps. We evaluate our methodology on CVC-ColonDB [8] and obtain detection results superior to the state-of-the-art [13].

2 Proposed Method

Our algorithms consists of two stages. In the first stage, the corresponding edge map for an input colonoscopy image is constructed and then refined through a set of boundary classifiers. In the second stage, our vote accumulation scheme is applied to the refined edge map and the polyp is localized.

2.1 Boundary Classification

We base our classification scheme on the observation that image appearance across the boundaries of polyps differ from the that of other boundaries in colonoscopy images. We propose to model this valuable source of information with a set of boundary classifiers. ***First***, we collect a crude set of edge pixels using Canny edge detector. ***Second***, we

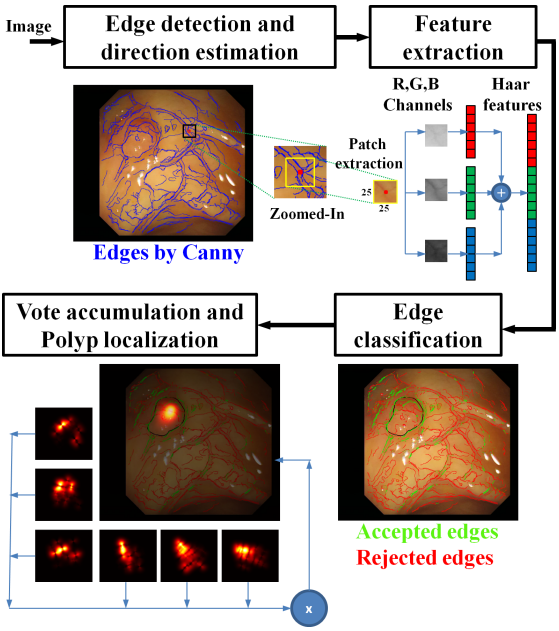


Fig. 2. The schematic overview of the proposed detection method. The black contour delineates the boundary of the polyp. The region with the highest values in the heat map is considered as the location of the polyp.

select an image patch around each edge pixel and then group the image patches into six categories according to the orientation of the central edge pixel of the patches. ***Third***, we train a classifier for the patches inside each category. The goal is to classify image patches into polyp and non-polyp categories, where the polyp category contains the patches whose central edge pixels lie on polyp boundaries, and the non-polyp category contains the patches whose central edge pixels are found on vessels, folds, wrinkles and other objects with strong boundaries in colonoscopy images. We use Haar features for training the classifiers. ***Given a test image***, edges are extracted and their corresponding patches are grouped into the 6 categories. Each category of patches is then fed to the corresponding classifier. Finally, edges whose corresponding image patches are assigned to the polyp category will be used in our vote accumulation scheme for polyp localization. Fig. 2 illustrates the polyp detection process for a test image.

Edge Detection and Edge Direction Estimation. A critical stage of our methodology is edge detection, since the desired edge pixels that are not captured by the edge detector cannot be recovered in the later stages. We apply Canny's method on three color channels to extract as many edges as possible. A fundamental parameter of the Canny is the standard deviation of Gaussian smoothing, σ_g, which as will be discussed in Section 3, if chosen in an appropriate range, will minimally impact the overall performance. We estimate edge direction by performing tensor voting [14] with the assumption that a ball tensor is placed at each edge pixel. In ball tensor voting, edge direction at a pixel is

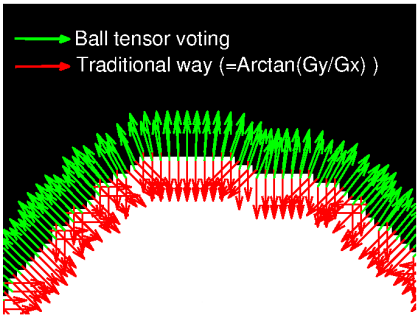


Fig. 3. Ball tensor voting determines edge directions more reliably than the traditional method based on horizontal and vertical image gradient

determined according to the arrangement of the surrounding edge pixels such that the continuation of edge direction is maintained—the locations of neighboring edges determine edge direction at a pixel. It is therefore very unlikely to obtain an inconsistent or non-smooth edge direction map. Fig. 3 compares the edge directions obtained with ball voting and the traditional way based on horizontal and vertical gradients. As seen, edge direction is more reliably estimated with ball tensor voting. A complete description of tensor voting is beyond the scope of this paper, interested readers are referred to [14] for more information.

Feature Extraction and Classification. We select a 25x25 image patch around each edge pixel and then extract Haar features from the R, G, and B channels of the collected patches. Haar features compute intensity differences between neighboring horizontal and vertical blocks in various scales and locations; a property that allows us to efficiently capture local image appearance across boundaries. One drawback of Haar features with regard to our particular application is that they are sensitive to the orientation of edge segments, meaning that different Haar patterns are required to capture image appearance across edge segments lying at different orientations. To overcome this drawback, we break the original classification task into six smaller problems by grouping the image patches into six categories according to the orientation of their central edge pixels, where each category covers $1/6$ of $[0, \pi]$: 0–30°, 31–60°, and so on. The patches inside each category exhibit less diversity compared to original problem, and thus the training process produces less complicated classifiers with more generalization power. For classification, we choose to use the random forest classifier [15] given its strong generalization power and its capability to avoid over-fitting of training data. Another distinguishing characteristic of the random forest classifier is the high quality probabilistic output [16] that we further utilize in the vote accumulation scheme. For the test stage, given a new image, image patches with classification confidence less than a threshold are discarded, meaning that their central edge pixels are excluded from the vote accumulation stage. Only edge pixels whose corresponding image patches pass the classification threshold will participate in the vote accumulation stage.

2.2 Polyp Localization

In the ideal classification scenario, all non-polyp edge pixels will be removed, and the arrangement of polyp edge pixels will indicate the location of polyps. However, in practice, a portion of non-polyp edges may pass the classification stage (false positives). On the knowledge that false positive edges often appear on elongated and low-curvature edge segments, we propose a vote accumulator that will mitigate the effect of false positive edges on polyp localization. The accumulation scheme assigns high values to the regions that are surrounded by curvy edges, but gives low values to the regions that are partially surrounded by elongated low curvature edges.

Voting Mechanism. Every positively classified edge pixel casts a scalar vote at each of its neighboring pixels (Fig. 4(a)). The magnitude of a vote cast by a voter pixel v at a receiver pixel $r = [x, y]$ is governed by

$$M_v(x, y) = C_v \times \exp(\frac{-\|v - r\|^2}{\sigma_F^2}) \times sin(\theta_{vr}), \tag{1}$$

where C_v is the probabilistic classification confidence assigned to the voter, the L^2-norm measures the Euclidean distance between the voter and receiver pixels, σ_F controls the size of voting field, and θ_{vr} is the angle formed by the voter, receiver, and edge orientation at the voter. The exponential and sinusoid decay functions (i) limit the contributions of distant edge segments, (ii) allow vote accumulation from the voters whose edge normals do not intersect at a certain point inside a polyp, and (iii) enable smooth vote propagation which will be later used in our ray back projection technique to determine the size of a polyp. According to this formula, pixels with smaller Euclidean distance to the voter and with larger acute angle θ_{vr} with respect to the edge direction receive votes with higher values.

The only parameter of our voting scheme is the size of voting field which is controlled by σ_F. Small values of σ_F makes the voting scheme sensitive to small regions with high vote accumulation, allowing for the detection of small polyps. On the contrary, large values of σ_F enables the detection of large polyps but it also allows for interference of farther voters. Since large polyps can be effortlessly detected by colonoscopists, σ_F should be adjusted for detecting polyps of small and moderate sizes. Considering that the missed polyps are usually 9 to 16 times smaller than the size of images, we allow σ_F to change between 70 and 90.

Vote Accumulation. We use the previous grouping criterion and divide the positively classified edges into six categories,$V_i|_{i=1}^{6}$. We then perform the voting process six times, at each time, we allow the edge pixels of one specific category to vote. Once all votes are cast, an accumulator adds up the votes received at each pixel and generates a voting map for each group of the edges. The resultant six voting maps are then multiplied to form the final map whose maximum indicates the location of a polyp candidate. Mathematically,

$$\underset{x,y}{\operatorname{argmax}} \prod_{i=1}^{6} \sum_{v \in V_i} M_v(x, y). \tag{2}$$

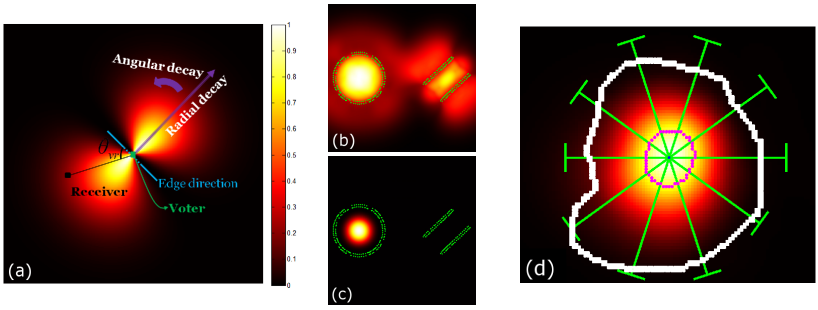


Fig. 4. (a) The geometric illustration of the voting scheme for an edge pixel lying at 135 degree. The voting map (b) without edge grouping and (c) with edge grouping for parallel and circular edges. As seen, the latter accumulator mitigates the undesirable vote accumulation between low curvature edge segments. (d) The voting map for a polyp (shown in white) and the determined search radii for a subset of radial rays.

The necessity of edge grouping prior to vote casting and accumulation is demonstrated in Fig. 4, where our voting scheme with and without edge grouping is compared for the edge pixels arranged on a circle and the edge pixels lying on two parallel lines. The resultant voting map with no edge grouping shows two regions with high vote accumulation: inside the circular edges which is desirable, and between the parallel lines which is undesirable (Fig. 4(b)). Preceding the voting scheme with edge grouping however mitigates the undesirable vote accumulation between low curvature edge segments and as shown in Fig. 4(c), only the region inside the circular edges receives high vote accumulation.

Ray Back-Projection. We propose a ray back-projection technique that measures the probability that a polyp candidate is a true positive detection. Such a probabilistic output is essential for rejecting false detections. The main idea is to cast rays from the detection point outward in all possible directions and then examine what fraction of rays hits the positively classified edges within a radius. The larger the fraction, the more likely a polyp candidate is a true detection.

The key to our ray back projection is how to determine the search radius for each individual ray. The radius must be large enough to include polyp boundaries and short enough to exclude false positive edge. Short or long radii may underestimate or overestimate the polyp likelihood. One way to estimate search radius is to examine decay pattern along each radial ray. In general, vote accumulation along each radial ray attenuates in two stages: (1) a slow Gaussian-like attenuation close to the detection point and (2) a more rapid attenuation as we distance from the high accumulation area. Our experiments reveal that if the decay had followed the slow Gaussian pattern, polyp boundary would have been reached within 3 standard deviations of the Gaussian function. We therefore aim to model a decay signal at angle θ with a Gaussian function and set the corresponding search radius to be $3\sigma_\theta$. Given two points on a decay signal $\{p_1, p_2\}$ and their corresponding vote accumulations $\{v_1, v_2\}$, one can obtain σ_θ as follows:

$$\sigma_\theta = \frac{\|p_2 - p_1\|^2}{-ln(v_2/v_1)}, \tag{3}$$

where p_1 contains image coordinates of the detection point and p_2 contains image coordinates of a point on the signal within 70% of maximum vote accumulation ($v_2/v_1 \geq 70\%$) , the range in which the decay signals exhibit a Gaussian-like decrease. In our experiments, we set $\frac{v_2}{v_1}$ to be 70%. Fig. 4(d) shows a polyp and its corresponding voting map. The dot points shows where the decay signal reach 70% of maximum vote accumulation and the green lines shows search radii for a subset of radial rays. Once search radii for all rays are determined, the probability of a true detection is measured as follows:

$$p(polyp|R_\theta|_{\theta=0}^{359}) = \frac{1}{180} \sum_{\theta=0:179} R_\theta \oplus R_{\theta+180}, \tag{4}$$

where R_θ is an indicator variable that takes 1 if the ray at angle θ hits at least 1 positive edge and 0 otherwise. This equation aims to treat both polyps with complete boundaries and those with partial boundaries that are located behind the folds or at image boarders equally, enabling us to detect partially appearing polyps.

Our voting scheme provides two major advantages over traditional Hough transform (HT). First, while HT is valued for detecting shapes with specific parametric model (e.g., circle and ellipse), our method naturally handles a variety of curvy shapes with local convex and concave boundaries. Second, HT does not produce a normalized output which complicates a classification threshold for accepting or rejecting polyp candidates; a limitation that is properly handled by the suggested ray back projection technique.

3 Experiments

We used *CVC-ColonDB* [8] database to evaluate our method. *CVC-ColonDB* is the only publicly available polyp database and consists of 380 colonoscopy images, which are selected from 15 short colonoscopy videos. This database contains 1 Pedunculated, 9 sessile, and 5 flat polyps. Bernal et al. collected *CVC-ColonDB* database and used 300 out of 380 colonoscopy images in their recent work [13] to evaluate their polyp appearance model. To have fair comparisons with [13], we also employed the same subset of images to test our methodology. In the following, we first evaluate each stage of the proposed polyp detection method and then provide the final results.

3.1 Edge Detection

Edge detection yields a crude set of candidate edges. The lower and upper thresholds of the Canny were computed automatically relative to the highest value of the gradient magnitude of the image. To determine the degree of Gaussian smoothing, σ_g, we performed a set of experiments and investigated how changes in Gaussian smoothing can affect the percentage of polyp edges that can be detected by the Canny in each of the 300 images. The result are shown in Fig. 5. Each box plot displays the distribution of polyp edge detection rates for the 300 images. As seen, the larger the σ_g, the lower the median of polyp edge detection rate. Since the polyp edges missed during the edge detection stage cannot be recovered in the later stages, we chose to set $\sigma_g = 3$ which achieves a high level of sensitivity against polyp edges. Fig. 5 also shows the fraction

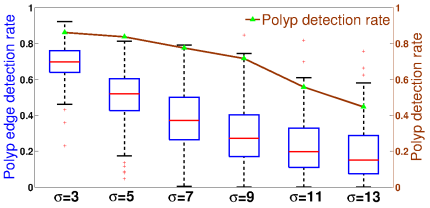


Fig. 5. Effect of Gaussian smoothing on the sensitivity of Canny edge detector and the sensitivity of voting scheme. Each box plot shows the percentages of polyp edges detected by Canny in 300 colonoscopy images.

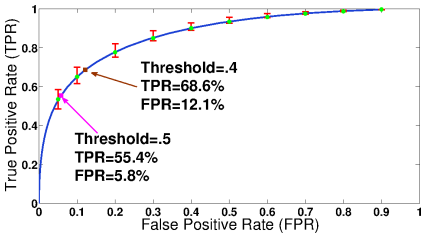


Fig. 6. The overall ROC curve of 5-fold cross validation. The blue plot is the combination of the ROCs obtained from the 5 test folds and the red bars show their range of variations. We use the default classification threshold.

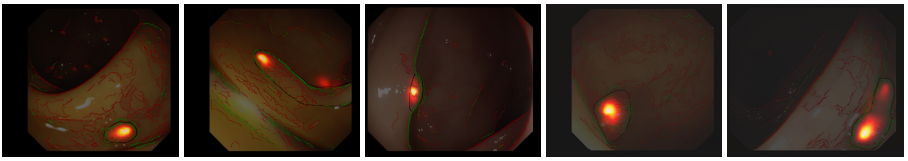


Fig. 7. Successful candidate casting. The green and red pixels indicate the accepted and rejected edges after classification. Voting maps produced by the accepted edges are superimposed on 5 colonoscopy images. As seen, the maximum of each voting map (a polyp candidate) falls inside the polyp region.

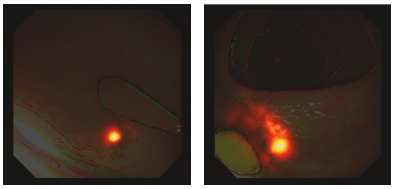


Fig. 8. Examples of unsuccessful candidate casting that are caused by (left) aggressive edge removal and (right) inadequate visible polyp boundary

Table 1. Precision and recall of our detection method and SA-DOVA[13] at 4 operating points. Limited by performance of the voting scheme, our detection method achieves up to 86% recall.

Recall	Precision	
	Our method	SA-DOVA [13]
50%	90%	92%
70%	89%	65%
86%	86%	60%
100%	—	50%

of images for which the maximum of voting map falls inside a polyp. As expected, polyp detection rate decreases as σ_g increases. Note that larger values of σ_g tend to exclude more polyp edges from the classification stage, leading to less accurate polyp localization.

3.2 Edge Classification

We employed 5-fold cross validation to use the totality of 300 images for performance evaluation. For each training set, we performed stratified sampling for selecting positive and negative patches. Such a sampling is required given the imbalanced nature of classification task and the very large number of polyp and non-polyp edges. The training image patches were then grouped into six categories and a random forest classifier was trained for each category of the patches. Next, the six classifiers were applied to the test images in order to classify all the edge pixels returned by the Canny. We obtained an ROC curve for all the test folds by putting together all the probabilistic outcomes generated by the six classifiers on the 5 test folds. In this way, we combined the 30 possible ROCs (6 ROCs for each of the 5 folds) into one single ROC plot as shown in Fig. 6. We used the default classification threshold to remove negative edges.

3.3 Voting Scheme

Our vote accumulator placed the maxima of voting maps inside 262 out of the 300 polyps which outperforms the state-of-the-art [13] where 252 polyps were detected by SA-DOVA descriptor. Fig. 7 shows examples of successful candidate casting where the maxima of voting scheme fall inside the polyp regions. The false candidates produced by our voting scheme mostly occurred due to aggressive edge classification or inadequate clear boundary between polyps and their surrounding area. Fig. 8 shows examples where the candidates were unsuccessfully placed outside the polyp regions.

3.4 Polyp Detection

We obtained precision and recall values by changing a threshold on the polyp likelihood and calculating the fraction of true and false detections at each threshold. We considered a detection as a "true detection" if the maximum of the voting map falls inside the ground truth contour provided in the database and the corresponding polyp likelihood passes the detection threshold. Since our polyp detection method placed the maxima of the voting maps in 262 out of 300 images, it achieved maximum recall of 86%. Table 1 shows precision and recall values at 4 operating points for our polyp detection algorithm and the one suggested in [13]. As seen, the obtained results are promising and outperform the state-of-the-art [13].

4 Conclusion

Our approach differs from the previous methods in systematically exploiting the unique appearance of polyp boundaries to suppress non-polyp edges, yielding cleaner edge maps for our novel vote accumulation, which can accommodate a large variation of polyp shapes, eliminate parallel edge configurations, and enable polyp detection from partially identified boundaries of polyps. Our polyp detection method requires only two parameters to be adjusted: σ_F, which is automatically set in a range given image/polyp size ratio, and σ_g, which gives comparable performance when changes between 3 and 7 (Fig. 5). As our future work, we plan to evaluate the suggested methodology on a significantly larger polyp database.

References

1. Siegel, R., Naishadham, D., Jemal, A.: Cancer statistics, 2013. CA: A Cancer Journal for Clinicians 63(1), 11–30 (2013)
2. Rabeneck, L., El-Serag, H., Davila, J., Sandler, R.: Outcomes of colorectal cancer in the united states: no change in survival (1986-1997). The American Journal of Gastroenterology 98(2), 471 (2003)
3. Winawer, S.J., Zauber, A.G., Ho, M.N., et al.: Prevention of colorectal cancer by colonoscopic polypectomy. New England Journal of Medicine 329(27), 1977–1981 (1993)
4. Heresbach, D., Barrioz, T., Lapalus, M.G., Coumaros, D., et al.: Miss rate for colorectal neoplastic polyps: a prospective multicenter study of back-to-back video colonoscopies. Endoscopy 40(4), 284–290 (2008)
5. van Rijn, J., Reitsma, J., Stoker, J., Bossuyt, P., van Deventer, S., Dekker, E.: Polyp miss rate determined by tandem colonoscopy: a systematic review. American Journal of Gastroenterology 101(2), 343–350 (2006)
6. Bressler, B., Paszat, L.F., Chen, Z., Rothwell, D.M., Vinden, C., Rabeneck, L.: Rates of new or missed colorectal cancers after colonoscopy and their risk factors: A population-based analysis. Gastroenterology 132(1), 96–102 (2007), http://www.sciencedirect.com/science/article/pii/S001650850602261X
7. Hewett, D.G., Kahi, C.J., Rex, D.K.: Does colonoscopy work? Journal of the National Comprehensive Cancer Network 8(1), 67–77 (2010)
8. CVC-Databasecolon: A database for assessment of polyp detection (2011), http://mv.cvc.uab.es/projects/colon-qa/cvccolondb
9. Karkanis, S., Iakovidis, D., Maroulis, D., Karras, D., Tzivras, M.: Computer-aided tumor detection in endoscopic video using color wavelet features. IEEE Transactions on Information Technology in Biomedicine 7(3), 141–152 (2003)
10. Park, S.Y., Sargent, D., Spofford, I., Vosburgh, K., A-Rahim, Y.: A colon video analysis framework for polyp detection. IEEE Transactions on Biomedical Engineering 59(5), 1408–1418 (2012)
11. Iakovidis, D.K., Maroulis, D.E., Karkanis, S.A.: An intelligent system for automatic detection of gastrointestinal adenomas in video endoscopy. Computers in Biology and Medicine 36(10), 1084–1103 (2006)
12. Hwang, S., Oh, J., Tavanapong, W., Wong, J., de Groen, P.: Polyp detection in colonoscopy video using elliptical shape feature. In: IEEE International Conference on Image Processing, ICIP 2007, vol. 2, pp. II-465–II-468 (2007)
13. Bernal, J., Sánchez, J., Vilariño, F.: Towards automatic polyp detection with a polyp appearance model. Pattern Recognition 45(9), 3166–3182 (2012)
14. Mordohai, P., Medioni, G.: Tensor Voting: A Perceptual Organization Approach to Computer Vision and Machine Learning. Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan & Claypool Publishers (2007), http://books.google.com/books?id=uvwxw5sJKywC
15. Breiman, L.: Random forests. Machine Learning 45(1), 5–32 (2001)
16. Criminisi, A., Shotton, J.: Decision Forests for Computer Vision and Medical Image Analysis. Springer (2013)

A Novel Computer Aided Detection (CADe) Scheme for Colonic Polyps Based on the Structure Decomposition

Huafeng Wang, Lihong Li, Hao Peng, Hao Han, Bowen Song, Yunhong Wang, Xianfeng Gu, and Zhengrong Liang

School of Software, Beihang University, Beijing, China, 100083
wanghuafengbuaa@gmail.com, lilong.li@csi.cuny.edu,
{happeng,haohan,bowen}@mil.sunysb.edu, yhwang@buaa.edu.cn,
gu@cs.sunysb.edu, Jerome.Liang@sunysb.edu

Abstract. Accurately detecting small polyps (ranged from 5~8mm) on the colon wall is of great significance for early diagnosis colorectal cancers. However, colon usually consists of the mucosa layers which result in partial volume effect (PVE) on the colon wall. Consequently, the task of computer aided detection (CADe) of polyps turns into too complicated to be reached by simply following solo philosophy. In order to achieve the mission of small polyps' detection, we propose a novel global structure decomposition approach in this paper. That is, the complex colon was separated into much uniform broken parts by means of analysis on second order derivatives of the volume image. Experimentally, we chose 60 patient cases from dataset provided by Wisconsin, and in which we focus on the polyps whose size range from 5~8mm to validate the presented new approach. Compared with previously presented in the literature, the experimental results are much more promising with an average sensitivity of 0.984. Meanwhile, the false positive rate dramatically decreased to 2.2 per dataset after false positive reduction.

Keywords: Colonic polyp, computed tomography colonography, computer-aided detection, colon structure decomposition.

1 Introduction

According to the American Cancer Society statistics, colorectal carcinoma is the third most commonly diagnosed cancer and the second leading cause of death from cancer in the United States [1]. It was estimated that 142,820 new cases will be diagnosed with 50,830 dying from the disease in 2013. Small polyps in the 6–9 mm range are typically pre-cancerous and are of interest for reporting during the exam; such polyps may be flagged for surveillance or surgical removal. In a preliminary report of polyp size measurements from the National computed tomography colonography (CTC) Trial, Chen et al. found that colonoscopy measurements were on average approximately 1.2 mm larger than those of CTC [2]. Hence, we will focus on polyps ranged from 5mm to 8mm polyps in our proposed method for computer aided detection (CADe) in CTC.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 63–72, 2013.

Based on the literature, the previous presented geometric based CADe method can mainly fall into two groups (surface-based and image-based) [3]. The first group (surface-based) is characterized of fitting locally a surface from the colon data, then computing surface curvatures from the fitted surface [4]. Unfortunately, because the actual colon wall is neither in a rigid shape nor with a single layer due to Partial Volume Effect (PVE) [5], it will be much difficult to obtain a continuous and smoothing representation of the colon inner wall in the case that we have no enough information of their complicated topology. In addition, usually the surface based method tend to smooth the data first, hence some details may be already lost at the first stage.

Fig. 1. The various growth morphology of small polyps under different background, where green arrow indicates the polyps' growing position

On the contrary, the volume based method put their emphasis on the calculation techniques of curvature for embedding iso-surface [6,7,8]. We conjecture that one of many factors contributing to the variation in detecting small polyps is the implicit assumption made for all the above CADe methods that the polyp is growing out from a flat surface environment. This assumption may be acceptable for larger polyps, but will be unlikely held for smaller ones less than 10 mm. This exploratory study aims to challenge this assumption and provides a reasonable theoretical description of more realistic polyp growing environment and the associated experimental outcome to demonstrate the gain of the more realistic polyp growing environment model (see Fig.1).

2 Methods

If we can distinguish different polyps' growing background in the colon lumen clearly, the subsequent CADe scheme will have high potential to achieve a better result. Thus, we propose a hierarchical or three-stage geometric analysis method in this paper. Fig. 2 Bellow shows the flowchart of our proposed CADe scheme, which can be dictated as follows. First, we performed interpolation of the original CTC dataset to construct isotropic CTC data. Then, we extracted the volumetric mucosa (VM) layer based o maximum a posteriori expectation maximization (MAP-EM) algorithm. After that, our CADe scheme went through three stages for polyps detection: 1) Geometric guided colon structure decomposition, 2) Shape feature based initial candidates detection, and 3) Knowledge based False Positive (FP) reduction.

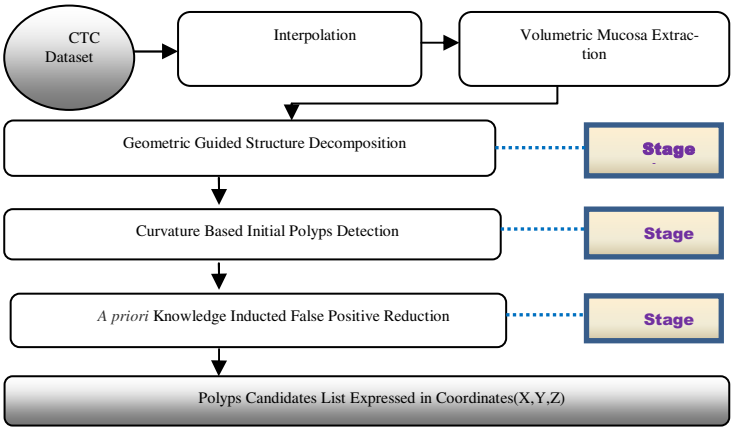
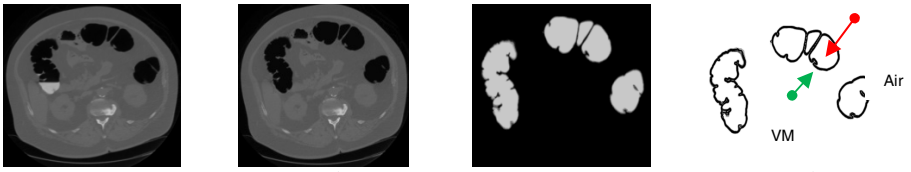


Fig. 2. An illustration of newly proposed CADe scheme

2.1 Volumetric Mucosa Extraction

To address the PVE, a MAP-EM algorithm which belongs to the soft image segmentation algorithm was proposed by Z. Liang et al [5]. The algorithm is capable of estimating the tissue mixture percentages inside each voxel. The output of the MAP-EM approach includes: (1) an electronically cleansed image in CT densities (Fig. 3(b)); and (2) the colon lumen (air class) percentage distribution map in the range of [0, 1], to reflect the PVE (Fig. 3(c)). For the simplicity of calculation, the value of each voxel in accordance with the existing percentage distribution is normalized to a value between 0 and 200.



(a) (b) (c) (d)

Fig. 3. MAP-EM segmentation results. (a) A slice of abdomen CT image with the tagged fecal material, (b) Electronic colon cleansing result, where the tagged material has been removed, (c) The segmentation result where the small intestine has been erased, (d) The corresponding VM regions (expressed as the dark area for a display convenience) with thickness where the further CADe will focus on.

2.2 Colon Decomposition

To build the priori information of the colon, we need first identify the structures of the VM region and name them with the certain labels, for example, planes, mountains, or

valleys. In general, given a 3D medical Image ***I(x,y,z),*** in order to calculate the geometric feature in this image, we need to calculate the Euclidean differential invariants(first or higher order derivatives). Usually derivatives are calculated as convolutions of the image with some predefined kernels or operators that try to approximate the derivatives (see Eq.1).

$$G_\sigma(x,y,z) = \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^3} e^{-\left(x^2+y^2+z^2\right)/2\sigma^2} \tag{1}$$

The σ determines the width of the Gaussian kernel. More specifically, it also tunes the filter response to the specific scale of lines. Let $2\sigma^2$= t be the scale-parameter, the process of generating a scale-space representation can be mathematically expressed as,

$$L(x,y,z;t) = G(x,y,z;t) * I(x,y,z) \tag{2}$$

We now compute the derivatives on the smoothed data. Consequently, we get the both first and second order derivatives of the smoothed volume image. And the second order Gaussian derivatives are further used to build a 3*3 Hessian matrix (Eq.3).

$$H = \begin{vmatrix} L_{xx} & L_{xy} & L_{xz} \\ L_{xy} & L_{yy} & L_{yz} \\ L_{xz} & L_{yz} & L_{zz} \end{vmatrix} \tag{3}$$

In order to acquire the shape apriority of colon volume by the above second order matrix, A. Frangi et al. suggested to calculate the eigenvalues$|\lambda_1| \le |\lambda_2| \le |\lambda_3|$ and the corresponding eigenvectors O, P and Q which form the orthogonal normal base for a local Cartesian coordinate system (Fig. 4)[9,10]. The vector O which is aligned to the line direction is the eigenvector with the smallest absolute Eigenvalue λ_1 .

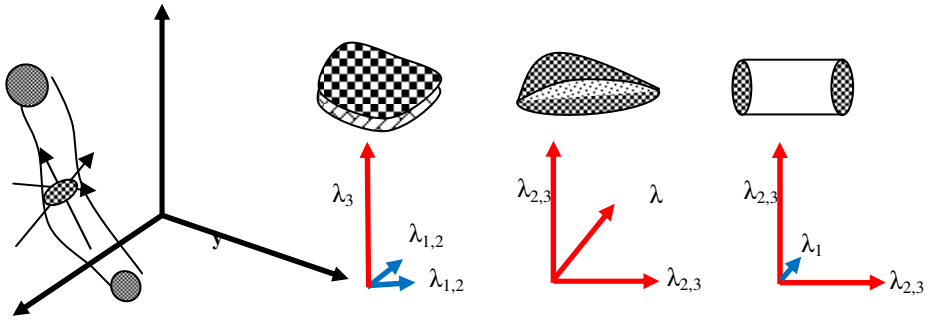


Fig. 4. Three kinds of typical colonic structure patterns: plate-like, blob-like and tube-like. We use different colors stand for various values: red stands for large values and blue stands for small values.

As mentioned above, the voxels of VM distributed in segmentation result have the normalized value $0 < VM_{Normalized\ Value} < 200$. Therefore, compared with the background value which was set by zero, they appear with a light gray value in the

image of segmentation. Thus, our goal is to decompose the light structures from each other within a dark background. In this respect, we have the following definitions,

We denotes the colon structural patterns on each voxel as P_i ,

$$P_i = \begin{cases} 50 & if(\lambda_1 < 0 \&\& \lambda_2 < 0 \&\& \lambda_3 < 0) \\ 100 & if(\lambda_1 > 0 \&\& \lambda_2 < 0 \&\& \lambda_3 < 0) \\ 150 & if(\lambda_1 > 0 \&\& \lambda_2 > 0 \&\& \lambda_3 < 0) \\ 0 & else \end{cases} \tag{4}$$

We further define the colon shape pattern measurement as the following,

$$T_t(\lambda) = \begin{cases} (1 - e^{-\frac{A^2}{2\sigma^2}}) \cdot (1 - e^{-\frac{S^2}{2\gamma^2}}) & if\ P_i == 50 \\ e^{-\frac{A^2}{2\sigma^2}} \cdot (1 - e^{-\frac{B^2}{2\beta^2}}) \cdot (1 - e^{-\frac{S^2}{2\gamma^2}}) & if\ P_i == 100 \\ e^{-\frac{B^2}{2\beta^2}} \cdot (1 - e^{-\frac{S^2}{2\gamma^2}}) & if P_i == 150 \\ 0 & otherwise \end{cases} \tag{5}$$

where,

$$A = \frac{\lambda_1}{\sqrt{|\lambda_2 \lambda_3|}}, B = \frac{|\lambda_2|}{|\lambda_3|}, S = \sqrt{{\lambda_1}^2 + {\lambda_2}^2 + {\lambda_3}^2} \quad , \tag{6}$$

and α, β, γ are user-defined parameters. The eigenvalues: $\lambda_1, \lambda_2, \lambda_3$ are with $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$. The $\boldsymbol{t}$ footer in T_i indicates that pattern likeness is calculated on a smoothed version of the image and is therefore a representative of the variations of image intensity at the spatial scale $\boldsymbol{t}$. Finally, the pattern likeness is estimated at a range of spatial scales, and the maximum response is selected at every voxel (see Eq.7). Hence, the algorithm becomes an adaptive one. In practice, α , β were fixed to 0.5. The value of the threshold γ depends on the grey-scale range of the image and half the value of the maximum Hessian norm has proven to work in most cases [11].

$$T_i(\lambda) = \max_{t \in [t_{min}, t_{max}]} T_t(\lambda) \ , \tag{7}$$

where $t_{min}\ and\ t_{max}$ are the maximum and minimum scales at which relevant structures are expected to be sorted out.

According to the pattern likeness value $\boldsymbol{T_i}$ at every voxel, the whole VM will be described in tube-like (mostly haustral wall), blob-like (mostly haustral folds or polyps) and other pattern likeness (mostly the connected or jointed areas with $T_i = 0$) (as shown in Fig.5).

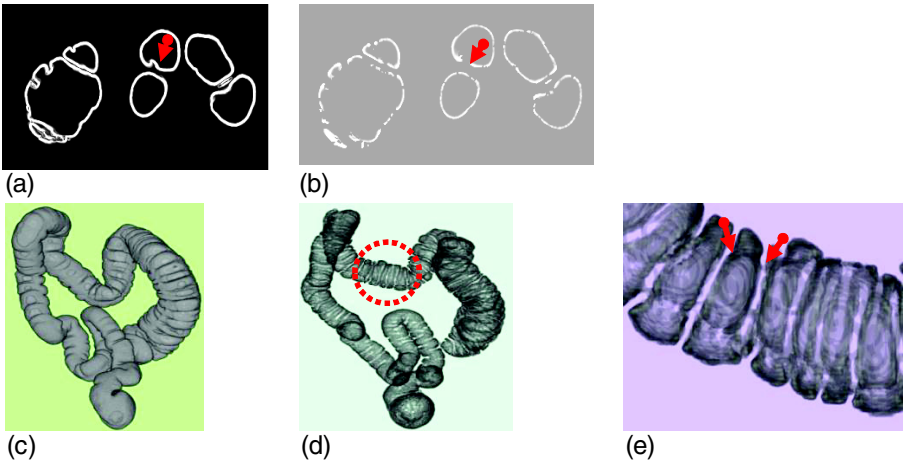


Fig. 5. 2D results of the measurements (first row) and 3D results by volume rendering (second row). Red arrow indicates the region of interest for demonstrating the enhancement of the colon wall.

2.3 Shape Feature Based Initial Candidate Detection

Given a smoothed volume data, the gradient g of the iso-surfaces can be determined by,

$$g = \nabla L = \left[\partial_x L \;\; \partial_y L \;\; \partial_z L\right]^T \quad , \tag{8}$$

where the definition of $\partial_x L$, $\partial_y L$ *and* $\partial_z L$ can refer to Eq.2.

Then the iso-surface normal n is defined as by assuming that the values of I increase from colon wall towards colon lumen,

$$n = -g/|g| \tag{9}$$

Curvature information will be contained in k_T, where k_T is calculated by,

$$k_T = \nabla \cdot n^T = -\frac{1}{|g|}(I - nn^T)H \quad , \tag{10}$$

where $\boldsymbol{I}$ is the intensity matrix, and H is the above mentioned Hessian matrix. The resulted principal curvatures and directions were expressed as functions of the first and second derivatives of the image. We can further calculate the shape index (SI) and curvedness (CV) by substituting the principal curvatures (k_1, k_2) for the new ones (K_{L1}, K_{L2}).

$$K_{L1,2}(x_i, y_i, z_i) = \frac{\Sigma_j^{\vec{p}} W_j k_T}{\Sigma_j^{\vec{p}} W_j} \quad , \tag{11}$$

where $\vec{p}$ is the path along which the local curvature is integrated. And W_j is the weight which is determined by the relative distance between point (x_i, y_i, z_i) and all the points on the integration path(see Fig.6).

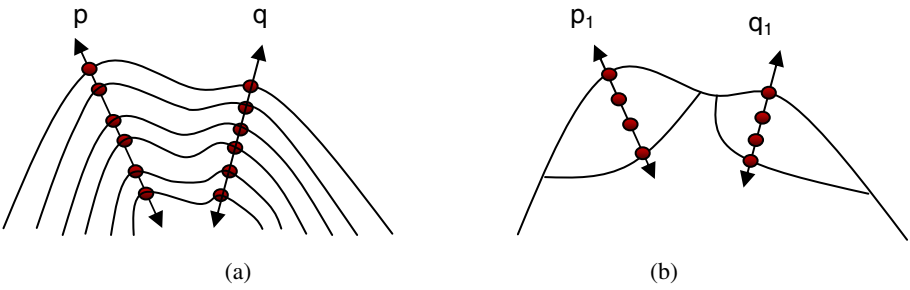


Fig. 6. The illustration of the integral curvature calculated on (a) the original mucosa layer and (b) the thinned mucosa layer

2.4 False Positive Reduction

As claimed by many previous researchers, the upper part of the colonic polyps usually appears as a "spherical cap" bulging into the colon lumen [8].Given SI and CV as obtained above, the variations on the colon wall with local "spherical cup" shapes can be detected by using of a simple clustering rule. However, the shape index and curvedness are much sensitive to small variations on the colon wall. Consequently, more than one dozens suspicious patches would be generated including a great number of non-polyp ones. To eliminate the non-polyp patches we need perform FP reduction.

Regarding the post-process of polyp candidates' selection, the quantitative of a priori knowledge, such as SI, CV, texture, polyp growing environments, and radial length, plays a very important role. Therefore, distinguishing the respective geometric structure of the colon is one of the key points to find true polyps from the initial candidates.

During the CADe, several geometric features can be extracted, e.g., SI, CV, volume and axis ratio, coverage ratio, and radiation ratio. These geometrical features will benefit the FP reduction process. Zhu et al. [6] claimed that it's difficult to identify the colonic structures from polyps by using only geometric information and they suggested analyzing the internal texture features for FP reduction. Similar to the textural analysis approach, we introduce an object likeness distribution analysis (OLDA) in our new FP reduction stage. For the purpose of conducting OLDA analysis, a volume of interest (VOI) is extracted for each suspicious patch. Because the object likeness on each voxel has been extracted already, the OLDA is measured for each patch as

$$OLDA_{mean} = \frac{1}{n}\sum_{i=1}^{n} T_i \quad and\ OLDA_{stdv} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(T_i - OLDA_{mean})^2} \qquad (12)$$

where $OLDA_{mean}$ is the mean of likeness value T_i , $OLDA_{stdv}$ is the standard deviation, and n is the total number of voxels contained in a suspicious patch.

To combine the OLDA with other features which contain the mean and variance of SI, the variance of CV, and the variance of CT intensity [12], a feature vector for each

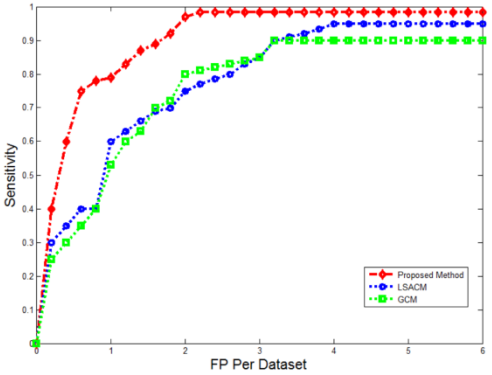


Fig. 7. FROC curve of the comparison based selected datasets

patch was built. Then the generated feature vectors were fed into the nonlinear RBF-SVM (radial basis function based support vector machine) [13]. In our study, we exploited the LibSVM package which was proposed by C. Chang et al [14].

3 Experiment and Result

We selected a CTC database of 60 patients with 120 CT scans from both supine and prone positions from the Wisconsin hospital. The patient's Age varies from 41 to 85 ; total number of the female patients is 13 and 57 is the male patients number. One polyp in different scans was counted as different polyps, and there were 130 clinically significant polyps (larger than or equal to 5 mm) confirmed by both OC and VC. Polyps smaller than 5 mm or larger than 8mm were not considered in this study.

Receiver Operating Characteristics (ROC) graphs are a useful and clear possibility for organizing classifiers and visualizing their performance, and the FROC (free-response ROC) curve is a generalization. It is defined as a plot of TP (true positive) number/total number of the polyps vs. FP(false positive) number/total number of the datasets. Note that the FP number here stands for the number after the FP reduction.

Compared to the GCM [8] which exploited typical volume based method and LSAC [15] which belongs to enhanced volume based shape analysis method, the new scheme significantly reduced the number of FPs. For our method the FROC curve analysis yields 2.2 FPs per dataset at 98.4% sensitivity (shown in Fig.7).

4 Discussion and Conclusion

In this study, we went thoroughly the challenges related to the small colonic polyps detection with both theoretical and experimental analysis. The new proposed method aims intuitively to make the much complicated structures be decomposed into simpler ones is of great advantages. In order to confirm if our new idea works very well in the much complicated situation of colon lumen, we reviewed both the successful cases

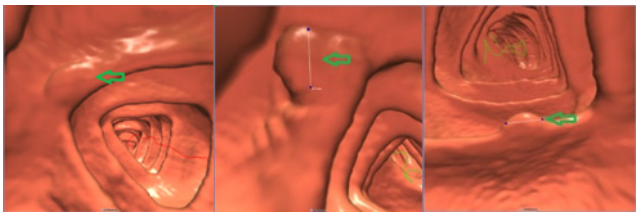


Fig. 8. The polyps overlapping with the haustral folds (indicated by green arrow)

and the failed cases. Among the successful cases, a few of them were previously not found by other methods, for example GCM or LSACM. Consequently, a very encouraging phenomenon was observed: the cases where the small polyps missed before but got by new proposed method are mostly those cases which are overlapping with haustral folds (Fig.8).

However, when we went through the failed cases (only two), one of the frequently failed reasons can be attributed to the over cleansing problem brought at the segmentation stage (as shown in Fig.9).

For the future work, we will focus on extremely eliminating the effects caused by infiltration of tagged material during the segmentation stage.

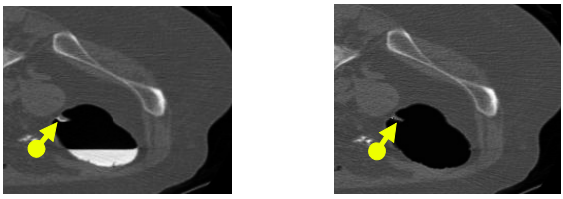


Fig. 9. The over-cleansing problem, where yellow arrow indicates the region of over cleansed

Acknowledgments. This work was partially supported by the NIH/NCI under Grant #CA143111, #CA082402, and the PSC-CUNY award #65230-00 43.

References

1. American Cancer Society: Cancer Facts & Figures. American Cancer Society, Atlanta (2013)
2. Chen, M.H., Blevins, M., Herman, B., Johnson, C.: National CT Colonography Trial (ACRIN 6664): are CT colonography, colonoscopy and pathology giving the same measures for polyp size? Radiological Society of North America, 279 (2008)
3. Liu, J., Kabadi, S., Uitert, R.V., Petrick, N., Deriche, R., Summers, R.M.: Improved computer-aided detection of small polyps in CT colonography using interpolation for curvature estimation. Med. Phys. 38, 4276 (2011)
4. Sundaram, P., Zomorodian, A., Beaulieu, C., Napel, S.: Colon Polyp Detection using Smoothed Shape Operators: Preliminary Results. Med. Image Anal. 12, 99–119 (2008)
5. Liang, Z., Wang, S.: An EM approach to MAP solution of segmenting tissue mixtures: A numerical analysis. IEEE Trans. Med. Imag. 28, 297–310 (2009)

6. Zhu, H., Liang, Z., Pickhardt, P.J.: Increasing computer-aided detection specificity by projection features for CT colonography. Med. Phys. 37, 1468–1481 (2010)
7. Liu, J., Kabadi, S., Uitert, R.V., Petrick, N., Deriche, R., Summers, R.M.: Improved computer-aided detection of small polyps in CT colonography using interpolation for curvature estimation. Med. Phys. 38, 4276 (2011)
8. Wang, Z., Liang, Z., Li, L.: Reduction of false positives by internal features for polyp detection in CT-based virtual colonoscopy. Med. Phys. 32, 3602–3616 (2005)
9. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Wells, W.M., Colchester, A., Delp, S. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)
10. Lorenz, C., Carlsen, I.C., Buzug, T.M., Fassnacht, C., Weese, J.: Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images. In: Troccaz, J., Mösges, R., Grimson, W.E.L. (eds.) CVRMed-MRCAS 1997. LNCS, vol. 1205, pp. 233–242. Springer, Heidelberg (1997)
11. Kim, S.H., Lee, J.M., Lee, J.G., Kim, J.H., Lefere, P.A., Han, J.K., Choi, B.I.: Computer-Aided Detection of Colonic Polyps at CT Colonography Using a Hessian Matrix-Based Algorithm: Preliminary Study. Gastrointestinal Imaging 189, 41–51 (2007)
12. Nappi, J., Frimmel, H., Dachman, A., Yoshida, H.: Computerized detection of colorectal masses in CT colonography based on fuzzy merging and wall-thickening analysis. Med. Phys. 31, 860–872 (2004)
13. Chang, Y.W., Hsieh, C.J., Chang, K.W., Ringgaard, M., Lin, C.J.: Training and testing low-degree polynomial data mappings via linear SVM. Machine Learning Research 11, 1471–1490 (2010)
14. Chang, C.-C., Lin, C.: LIBSVM: A library for support vector machines (2001), http://www.csie.ntu.edu.tw/~cjlin/libsvm
15. Zhu, H., Duan, C., Wang, S., Liang, Z.: Computer-aided detection of colonic polyps with level set based adaptive convolution. Cancer Management and Research 1, 1–13 (2009)

Computer-Aided Detection of Colorectal Lesions with Super-Resolution CT Colonography: Pilot Evaluation

Janne J. Näppi, Synho Do, and Hiroyuki Yoshida

Massachusetts General Hospital Imaging and Harvard Medical School,
25 New Chardon Street, Suite 400C, Boston, Massachusetts 02114, USA
jnappi@partners.org, sdo@nmr.mgh.harvard.edu,
yoshida.hiro@mgh.harvard.edu

Abstract. Reliable computer-aided detection (CADe) of small polyps and flat lesions is limited by the relatively low image resolution of computed tomographic colonography (CTC). We developed a sinogram-based super-resolution (SR) method to enhance the images of lesion candidates detected by CADe. First, CADe is used to detect lesion candidates at high sensitivity from conventional CTC images. Next, the signal patterns of the lesion candidates are enhanced in sinogram domain by use of non-uniform compressive sampling and iterative reconstruction to produce SR images of the lesion candidates. For pilot evaluation, an anthropomorphic phantom including simulated lesions was filled partially with fecal tagging and scanned by use of a CT scanner. A fully automated CADe scheme was used to detect lesion candidates in the images reconstructed at conventional 0.61-mm and at 0.10-mm SR image resolution. The proof-of-concept results indicate that the SR method has potential to reduce the number of FP CADe detections below that obtainable with the conventional CTC imaging technology.

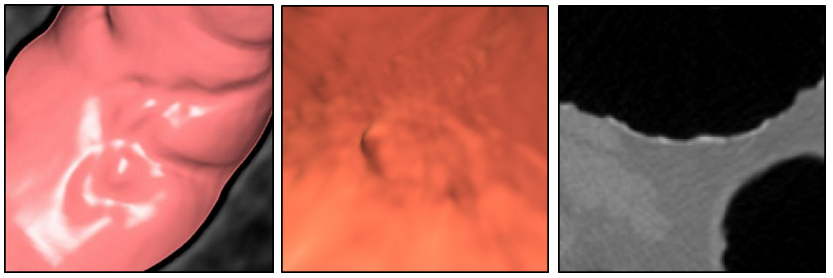
Keywords: Iterative reconstruction, super-resolution, domain decomposition, computer-aided detection, virtual colonoscopy.

1 Introduction

Although computed tomographic colonography (CTC) is able to detect large advanced polyps at high sensitivity [1], the detection of small polyps 6 – 9 mm and flat lesions has been shown to be more challenging [2]. In principle, computer-aided detection (CADe) systems may be able to detect such lesions at high sensitivity, but the presence of poorly tagged stool and thin layers of fecal tagging that imitate the shape and radiodensity of lesions makes it challenging to differentiate the detections of true lesions reliably from false positives (Fig. 1) [3]. Therefore, even if CADe can yield high detection sensitivity, the number of associated FP CADe detections is too high for routine clinical practice.

In this study, we considered the reconstruction of super-resolution (SR) images for improving the conspicuity of subtle lesions and for identifying partial-volume

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 73–80, 2013.



(a) (b) (c)

Fig. 1. Virtual endoluminal CTC views of (a) a true 20-mm flat lesion and (b) a thin layer of fecal tagging imitating a flat lesion. (c) Axial CTC image of a thin layer of fecal tagging on colon surface.

artifacts that imitate lesions. The enhanced detail provided by the SR images could be used to improve the detection accuracy of CADe beyond that obtainable with conventional CTC.

2 Methods

2.1 Super-Resolution CTC (SR-CTC)

To calculate SR images, we use model-based iterative CT image reconstruction (IRT) [4]. The reconstruction problem can be formulated as

$$\hat{f} = \text{argmin}_f E_f(g, f) + \alpha E_r(f), \tag{1}$$

where $E_f(g, f)$ is a data fidelity term between image f and sinogram g, $E_r(f)$ is a regularization term, and α is the weighting term. The data fidelity term can be formulated as

$$E_f(g, f) = ||g - Hf||^2 = \sum_j |g_j - Hf|^2, \tag{2}$$

where H is the system matrix. The general form of the regularization term can be formulated as the L_p norm

$$E_r(f) = ||Lf||^p = \sum_i |Lf_i|^p. \tag{3}$$

With $L = \nabla$ and $p = 1$, Eq. (3) becomes a total-variation regularizer that suppresses noise and preserves edges in the reconstructed images. To minimize the energy functional of Eq. (1), a lagged diffusivity fixed-point method is used for iterative approximation of the cost by a weighted quadratic cost and by solving for the resulting linear normal equations using pre-conditioned conjugated gradient iterations [5].

To reconstruct SR images, we combine Eq. (1) with fractional super-sampling of the sinogram space, where each simulated ray for a single detector is over-sampled with a bundle of casted virtual rays. By collection of the variables into a vector-matrix equation, the SR images can be calculated as the solution of Eq. (1).

The IRT method is computationally demanding, but the computation time is reduced by parallelization of the ray projections and by use of a fast non-iterative version of Siddon's radiological ray-tracing algorithm [6].

2.2 Computer-Aided Detection (CADe) System

Figure 2 presents an overview of the steps of the fully automated CADe system. The input CTC images are interpolated to isotropic resolution, and a thick region encompassing the region of colon surface is extracted from the CTC images [7]. Lesion candidates are detected by use of volumetric shape features [8]. Level-set segmentation is used for exctracting the regions of detected lesion candidates [9]. For false-positive (FP) reduction, several advanced texture and shape features are calculated from the regions of lesion candidates [10]. Based on the features, a random-forest classifier is used to calculate the lesion-likelihood of each lesion candidate. The lesion candidates that yield the highest likelihood represent the final output of the CADe system [11].

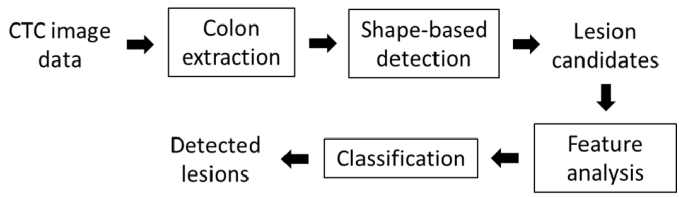


Fig. 2. Overview of the computation steps of the CADe system. A set of initially detected lesion candidates is identified from the input CTC data, after which the final detected lesions are determined by use of machine learning methods.

2.3 Application of SR-CTC to CADe

A complete SR image reconstruction of the input CTC image data would be computationally demanding and would require large amounts of memory. Because the CADe system can already detect 99 – 100% of all visible lesions on conventional CTC images (with a large number of FP detections) [12], we can reduce the computational demands of the SR-CTC scheme substantially by performing iterative reconstruction only for regions of interest (ROIs) that contain the initially detected lesion candidates. The other regions of CTC images need not to be considered.

Figure 3 shows an overview of the proposed detection scheme. First, the CTC images are reconstructed from original sinograms by use of the filtered back-projection (FBP) method at conventional CTC image resolution, and the CADe is used to detect the regions of initial lesion candidates from the reconstructed images.

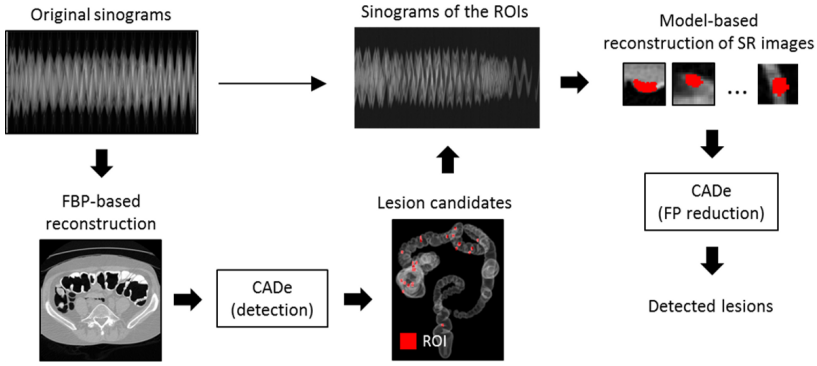


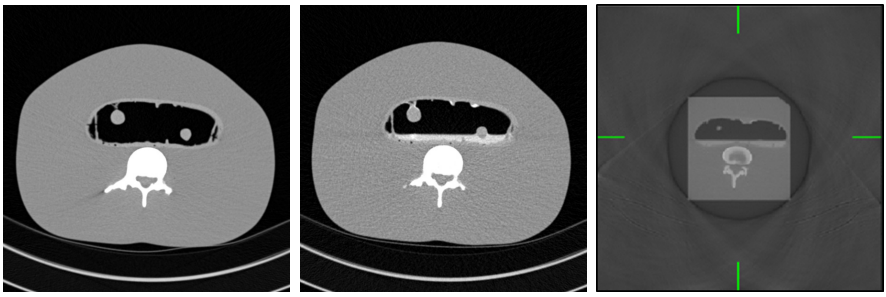
Fig. 3. Overview of the major steps of the SR-CTC detection scheme

Next, we use a projection-driven technique to reduce the reconstruction problem to a set of ROIs that encompass the detected lesion candidates and their immediate neighborhood region [13]. In the technique, the ROIs are reprojected to the sinogram domain. By nullification of the locations of ROIs, we create a set of projections that correspond to the relevant ray-projection field excluding the ROIs. By subtraction of this background region from the data, we facilitate iterative reconstruction of the ROIs at SR by the methods of Section 2.1. The technique makes the SR computations orders of magnitude smaller than what would be required by a full algebraic reconstruction.

Finally, the reconstructed SR images are reviewed by the CADe scheme to reduce FP detections. The output of this step determines the output of the CADe scheme.

3 Evaluation

An anthropomorphic human-colon phantom (The Phantom Laboratory Inc., Salem, NY) was designed to imitate features observed in clinical human CTC scans (Fig. 4a). The phantom was filled partially with an iodine-based tagging agent (Oxilan, Guerbet, Bloomington, IN) and with cereal to imitate residual materials (Fig. 4b). The phantom was scanned in supine position with a dual-energy CT scanner (SOMATOM Definition Flash, Siemens Healthcare) at tube currents of 22 mA (140 kVp) and 106 mA (80 Kvp). The CT scanner reconstructed the conventional CT images of the scans by use of a conventional FBP algorithm at a 0.61-mm image resolution. Only the 140 kVp scan was used for the reconstructions — dual-energy reconstructions were not considered.



(a) (b) (c)

Fig. 4. (a) An axial image of the anthropomorphical phantom. (b) For experiments, the anthropomorphic phantom was filled partially with a tagging agent and scanned at low dose. (c) Axial image of the region of transverse colon reconstructed at SR by use of the proposed projection-driven technique.

For pilot evaluation of the proposed ROI and SR methods, the raw sinogram data was downloaded from the CT scanner. To simplify the experiments, we reconstructed the complete region of transverse colon from the sinogram data at SR (Fig. 4c). The locations of lesion candidates that the CADe system detected on conventional CTC images were mapped to the SR-CTC images.

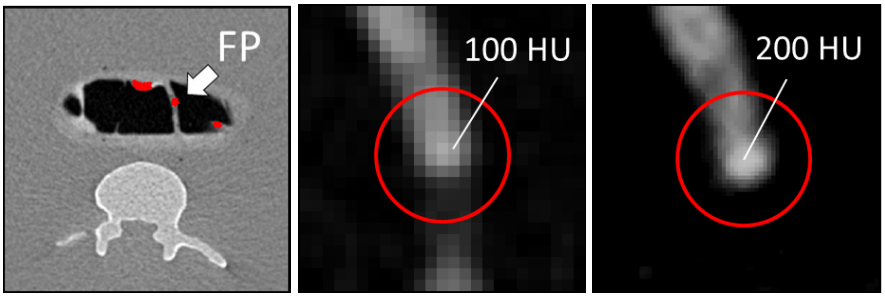
4 Results

The reconstructed region of transverse colon contained 6 simulated lesions. There were 2 pedunculated polyps, 2 sessile polyps, and 2 flat lesions. They measured between 8 – 16 mm in largest diameter. All 6 lesions were detected correctly by the CADe system on conventional CTC images. The use of SR did not cause CADe to reject any of the true positives.

The CADe system detected one FP lesion candidate on conventional CTC images (Fig. 5a). Magnification of the detected region at conventional image resolution (Fig. 5b) indicates that the detection could be a small polyp on a fold. The relatively low CT value of the detection on conventional CTC images (≈100 Hounsfield units (HU)) further suggests that the detection may be a soft-tissue lesion.

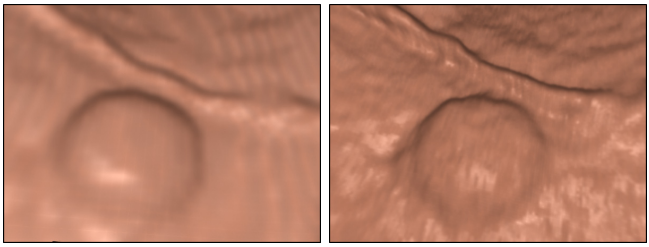
However, on the corresponding SR image (Fig. 5c), the detection has a relatively high radiodensity (200 HU). This indicates that the detection is caused by a small round droplet of fecal tagging on fold. Therefore, the use of SR-CT enabled the CADe system to exclude the lesion candidate confidently as a FP detection.

Figure 6 demonstrates the difference between conventional and SR images in the virtual endoscopic rendering of a simulated 8-mm polyp. Same rendering parameters were used in both cases. It is evident that the SR-CTC image shows the bowel surface in more detail than the conventional CTC image.



(a) (b) (c)

Fig. 5. (a) Axial CT image of CADe detections (red color) on the phantom. The arrow indicates a FP CADe detection on a fold. (b) Magnification of the indicated FP CADe detection (red circle) at a conventional 0.61-mm CTC image resolution. (c) Super-resolution image of the indicated FP CADe detection at a 0.1-mm image resolution.



(a) (b)

Fig. 6. Virtual endoscopic view of a simulated 8-mm polyp. (a) Rendered at conventional 0.61-mm image resolution. (b) Rendered at 0.1-mm super-resolution.

5 Discussion

The limited image resolution of conventional CTC makes it challenging to differentiate subtle lesions reliably from tagged feces and their partial-volume effects. This presents a potential limit to the detection accuracy of conventional CADe. However, enhancement of the lesion candidates by use of SR-CTC could be used to overcome this limitation.

The ROI reconstruction problem that was considered in this study is different from ROI tomography. In ROI tomography, the reconstruction is based on a local projection of the target region, whereas in this study, we assumed access to the entire projection data. Thus, the proposed method presents a practical post-acquisition approach for ROI-focused CTC.

The proposed SR method makes it possible to reconstruct images of detected ROIs at a resolution that is beyond those of the nominal resolution of the CT detector and the capabilities of conventional image reconstruction provided by the CT scanner. Thus, our approach is different from providing detected locations back to the CT scanner for performing conventional image reconstruction.

The SR-CTC may also have other uses than improvement of the standalone detection accuracy of CADe. Several studies have shown that radiologists can incorrectly dismiss correct CADe prompts of true lesions [14]. By providing enhanced SR-CTC images of the CADe detections, radiologists might be able to use the CADe prompts more effectively.

In this study, we reconstructed SR images of ROIs by use of IRT. The IRT has also several other advantages over the conventional FBP method. It can be used to reconstruct CT images at a substantially lower radiation dose than that of conventional standard-dose FBP, without impairing image quality. It can also be used to correct for beam-hardening artifacts that may appear at low radiation doses or with high concentrations of fecal-tagging agents [15].

This pilot study had several limitations. First, the proposed method was tested on an anthropomorphic phantom only. Second, the evaluation was largely anecdotal, because all true lesions were detected and there was only one false positive to be excluded by the use of SR. Therefore, it is premature to perform a formal receiver-operating characteristic (ROC) analysis for evaluating the CADe system performance and for comparing with expert reading.

Acknowledgements. This work was supported in part by grants of CA095279, CA166816, and CA131718.

References

1. 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)
2. Park, S., Kim, S., Lee, S., Bogoni, L., et al.: Sensitivity of CT colonography for nonpolypoid colorectal lesions interpreted by human readers and with computer-aided detection. Am. J. Roentgenol. 193, 70–78 (2009)
3. Näppi, J., Yoshida, H.: Virtual tagging for laxative-free CT colonography: pilot evaluation. Med. Phys. 36, 1830–1838 (2009)
4. Do, S., Cho, S., Karl, W., Kalra, M., Brady, T., Pien, H.: Accurate model-based high resolution cardiac image reconstruction in dual source CT. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 330–333. IEEE Press, Boston (2009)
5. Do, S., Karl, W., Liang, Z., Kalra, M., Brady, T., Pien, H.: A decomposition-based CT reconstruction formulation for reducing blooming artifacts. Phys. Med. Biol. 56, 7109–7125 (2011)
6. Jacobs, F., Sundermann, E., De Sutter, B., Christiaens, M., Lemahieu, I.: A fast algorithm to calculate the extract radiological path through a pixel or voxel space. J. Comput. Inf. T. 6, 89–94 (1998)
7. Näppi, J., Yoshida, H.: Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography. Acad. Radiol. 25, 287–300 (2007)
8. Yoshida, H., Näppi, J.: Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. IEEE Trans. Med. Imaging 20, 1261–1274 (2001)
9. Näppi, J., Frimmel, H., Dachman, A., Yoshida, H.: Computerized detection of colorectal masses in CT colonography based on fuzzy merging and wall-thickening analysis. Med. Phys. 31, 860–872 (2004)

10. 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)
11. 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)
12. Näppi, J., Yoshida, H.: Ranking of polyp candidates for CAD in CT colonography. In: SPIE Medical Imaging 2005: Physiology, Funcions, and Structure from Medical Images, vol. 5746, pp. 432–439 (2005)
13. Do, S., Liang, Z., Karl, W., Brady, T., Pien, H.: A projection-driven pre-correction technique for iterative reconstruction of helical cone-beam cardiac CT images. In: Proc. SPIE Medical Imaging: Physics of Medical Imaging, vol. 6913. SPIE Press (2008)
14. Taylor, S., Robinson, C., Boone, D., Honeyfield, L., Halligan, S.: Polyp characteristics correctly annotated by computer-aided detection software but ignored by reporting radiologists during CT colonography. Radiology 253, 715–723 (2009)
15. Hsieh, J., Molthen, R., Dawson, C., Johnson, R.: An iterative approach to the beam hardening correction in cone beam CT. Med. Phys. 27, 23–29 (2000)

Computer-Aided Detection of Non-polypoid Flat Lesions in CT Colonography: Observer Performance Study

Yasuji Ryu, Janne J. Näppi, Minh Phan, 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
{yryu,jnappi,mphan4}@partners.org,
yoshida.hiro@mgh.harvard.edu

Abstract. To evaluate the effect of computer-aided detection (CADe) on the performance of human readers in the detection of non-polypoid flat lesions from a large computed tomography (CT) colonography population. A total of 153 cathartic CT colonography cases, including 45 colonoscopy-confirmed, morphologically flat lesions, were sampled from a European multi-center CT colonography trial for asymptomatic patients at increased risk of colorectal cancer. Two readers (expert and non-expert) reviewed the 153 CT colonography cases and recorded all detected lesions using primary 3D interpretation and a CADe second-read paradigm. There were 17 patients with 18 flat lesions ≥10 mm in size and 17 patients with 27 flat lesions 6 – 9 mm in size. For the flat lesions ≥10 mm, per-patient sensitivities of the expert reader for unassisted and CADe-assisted readings were 59% [95% CI: 36–78%] and 71% [47–87%], respectively, whereas those of the non-expert reader were 41% [21–65%] and 47% [37–59%], respectively. For 6-9 mm flat lesions, the corresponding per-patient sensitivities of the expert reader were 59% [36–78%] and 76% [53–89%], respectively, whereas those of the non-expert were 47% [37–59%] and 82% [59–93%]. The results indicate that the use of CADe can increase the sensitivity of human readers in the detection of flat lesions in a screening setting.

Keywords: CT colonography, flat lesions, non-polypoid lesions, computer-assisted detection, observer study.

1 Introduction

Colon cancer is the second leading cause of cancer deaths in the United States, where more than 50,000 people die each year from colon cancer [1]. A majority of colon cancers develops from large benign polyps which grow on the bowel mucosa. Early detection and removal of the polyps can reduce the likelihood of developing colon cancer [2].

Minimally invasive population screening for colorectal cancer could be implemented by use of computed tomographic (CT) colonography [3]. In the National CT Colonography Trial (ACRIN 6664) with 2,531 participants from 15 U.S. centers, CT colonography yielded a per-patient sensitivity of 90% for large adenomatous polyps

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 81–88, 2013.

≥10 mm in size, on a par with the sensitivity of optical colonoscopy (OC) [4]. Recently, CT colonography was also listed as a viable screening option for colorectal cancer in the joint guideline from the American Cancer Society, the U.S. Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology [5].

However, the performance of CT colonography for the detection of morphologically flat lesions has been under increasing scrutiny after a recent study [6] which reported that the prevalence of non-polypoid flat colorectal neoplasia in U.S. screening (n = 616), surveillance (n = 654), and symptomatic (n = 549) populations was as high as 5.84%, 15.44%, and 6.01%, respectively. This is a concern, because flat neoplasia were previously not believed to exist in U.S. populations, and because studies have indicated that flat neoplasia are associated frequently with in situ or submucosal invasive carcinomas [6].

On CT colonography, non-polypoid lesions appear as plaque-shaped mucosal elevations with or without a depression. The accuracy of CT colonography for flat lesions varies among studies and it is still controversial. Park et al. [7] reported that fewer than 50% of flat lesions could be visualized on CT colonography in a study of 213 patients with 32 flat polyps even when lesions that were missed due to technical failures were excluded. However, more recently, Pickhardt et al. [8] reported excellent results from a prospective single-center screening study of 5,107 patients: CT colonography detected 125 polyps ≥6 mm in size categorized prospectively as flat polyps, and there were no histologically advanced lesions among the nine false-negative flat lesions seen only on colonoscopy.

The sources of the varying accuracy are not known completely; however, two important factors that potentially contribute to the large variation are (1) radiologists' perceptual errors because of the subtlety of morphologically flat lesions, and (2) suboptimal CT techniques [9]. Such errors could be complemented by the use of computer-aided detection (CADe), which automatically detects colorectal lesions in CT colonography data and indicates their locations to improve radiologists' diagnostic performance [10-12]. Recent study by Regge et al reported [13] excellent results from a prospective multicenter trial—it showed that CAD as 2nd reader improved per-patient sensitivity from 66% to 73%. Other retrospective studies also showed improved sensitivity in CAD assisted read compared to CAD unassisted read [14-15]. To the best of our knowledge, no CADe schemes have been designed specifically for the detection of flat lesions, whereas conventional CADe schemes have reported only modest success in the detection of flat lesions. Taylor et al. [16] collected a consecutive series of 24 morphologically flat T1 colorectal cancers undergoing staging CT colonography after assessment for potential endoscopic removal. A CADe software for polyp detection (ColonCADe API 4.0, Medicsight plc) detected 83% and 54% of the 24 cancers at false-positive (FP) rates of 37 and 10 per patient, respectively [16]. The high sensitivity of 83% suggests that conventional CADe may be able to detect flat lesions, but the large number of FP detections suggests that current CADe systems will need to be adapted to the morphological characteristics of minimally elevated flat lesions to produce the high levels of specificity reported for the detection of polypoid lesions.

In this study, we evaluated the performance of two radiologists in the detection of flat lesions using a CADe system that was optimized for the detection of flat lesions

on CT colonography images. The evaluation was performed with 153 clinical CT colonography cases from a European multi-center CT colonography trial [18].

2 Methods

2.1 Materials

A total of 1103 individuals at increased risk of colon cancer were recruited at 11 Italian centers and 1 Belgian center, and 937 participants were included in the final analysis [18]. For the current retrospective study, 153 of the CT colonography cases were used, which consisted of 119 normal cases and 34 abnormal cases with at least one non-polypoid flat lesion that was retrospectively visible on CT colonography images. These abnormal cases were chosen so that most of the colonoscopy-confirmed non-polypoid flat lesions observed in the clinical trial were included in this study. The normal cases were randomly selected from the trials cohort so that the approximate prevalence of the colonic lesions (including non-polypoid and polypoid lesions) was close to the increased-risk patient cohort of 20%. The presence of the flat lesions was determined based on the final report from the clinical trial.

2.2 CT Colonography

Each participant underwent a CT colonography examination which was followed by same-day OC at the same center. No specific colon cleansing directions were given to the participating centers, except that full bowel purgation was required and internationally recognized quality standards had to be met [18]. Hydro-soluble iodine agents alone or in combination with barium sulfate were accepted for orally administered fecal tagging. The CT colonography was performed in supine and prone positions with 120 kVp, ≤50 mA effective current, and a section thickness ≤2.5 mm. Intravenous contrast medium was not used.

2.3 Review Process

An expert radiologist who did not otherwise participate in the study annotated the precise locations of flat lesions in the CT colonography data based on prospective CT colonography and segmentally unblinded OC results. According to the universally adopted segmental checking procedure, a lesion reported at CT colonography was matched to a corresponding one reported at OC when it was located in the same or adjacent colon segment and when its size differed by no more than 50%. The matching was performed immediately after the conclusion of both tests, if necessary, by reviewing colonoscopy video registration and CT colonography. The lesion size was measured at colonoscopy using open biopsy forceps.

2.4 CADe System

The CADe system that was used in this study was fully automated (including fully automated colon segmentation) [17]. The detection algorithm was developed with 203

CT colonography cases that are different from the evaluation materials of this study. The development cases represented a heterogeneous variety of cathartic, reduced-cathartic, and non-cathartic bowel preparations from several U.S., European, and Asian medical centers. They included 20 flat lesions ≥10 mm in size and 29 flat lesions 6 – 9 mm in size. A previous standalone evaluation of the un-optimized CADe system with 73 patients and 107 lesions had yielded 91% per-lesion sensitivity at a median of 6 FP prompts per patient, where the patients had been sampled from the same general population as those of this study.

The CADe algorithm was optimized to detect flat lesions in the development cases without redundant false positives. For colon segmentation, the thickness of the extracted target region was optimized to maximize detection sensitivity for flat lesions. The feature values of the volumetric shape features that are used for the initial detection of candidate sites were thresholded at an optimal value range to detect all flat lesions. The smallest volume of a detected candidate region was optimized to detect all flat lesions without redundant false positives.

For FP reduction, shape and texture features that yielded highest area under the receiver-operating characteristic curve and that had low correlation with the other features in the detection of flat lesions were determined. Using the optimized features, a statistical random-forest classifier was constructed to differentiate flat lesions from other detections by use of a 100-fold cross-validation protocol. The random forest that yielded highest sensitivity for advanced flat lesions (adenomas and carcinomas) with the smallest average number of FP detections per patient was chosen as the classifier that was included in the final CADe algorithm.

The CADe output was integrated into a CT colonography reading workstation (Virtual Place, Aze Ltd., Tokyo, Japan) by use of a plugin interface. In CADe mode, the workstation displays a list of automatically detected lesions, and the CADe prompts are visualized as circles and arrows centered at the spatial locations of detections on two-dimensional and three-dimensional CT colonography images, respectively.

2.5 Image Interpretation

The expert reader had more than 15 years of experience in the interpretation of abdominal CT and had interpreted >500 CTC examinations. The non-expert reader was a novice reader who received CTC reading training with 100 cases. Electronic cleansing was not employed in this study. The case reading order was designed to distribute positive CT colonography cases evenly across all of the cases. The per-patient sensitivity for flat lesions was compared between unassisted and CADe-assisted readings. Prior to the observer study, both readers received training involving the reading of 20 CT colonography cases using the primary 3D interpretation and CADe second-read paradigm. The training cases were otherwise not used in this study.

Table 1. Per-patient detection performance of the expert reader

	≥6 mm			6-9 mm			≥10 mm		
	w/o CAD	w/ CAD	p-value†	w/o CAD	w/ CAD	p-value†	w/o CAD	w/ CAD	p-value†
Sensitivity	59% (20/34)	74% (25/34)	0.07	59% (10/17)	76% (13/17)	0.25	59% (10/17)	71% (12/17)	0.48
Specificity	78% (93/119)	79% (94/119)	1						

3 Results

There were 34 patients (22%) with 45 morphologically flat lesions: 17 patients had 18 flat lesions ≥10 mm in size and 17 had 27 flat lesions 6-9 mm in size. The standalone CADe system yielded 94% per-patient (89% per-lesion) sensitivity for the flat lesions with an average number of four CADe prompts per patient. For polypoid (non-flat) lesions, the CADe system yielded 76% per-patient (78% per-lesion) sensitivity. Table 1 and Table 2 show a summary of the per-patient reading results of the expert reader and the non-expert reader, respectively.

Table 2. Per-patient detection performance of the non-expert reader

	≥6 mm			6-9 mm			≥10 mm		
	w/o CAD	w/ CAD	p-value†	w/o CAD	w/ CAD	p-value†	w/o CAD	w/ CAD	p-value†
Sensitivity	44% (15/34)	65% (22/34)	0.02	47% (8/17)	82% (14/17)	0.04	41% (7/17)	47% (8/17)	1
Specificity	92% (110/119)	88% (105/119)	0.07						

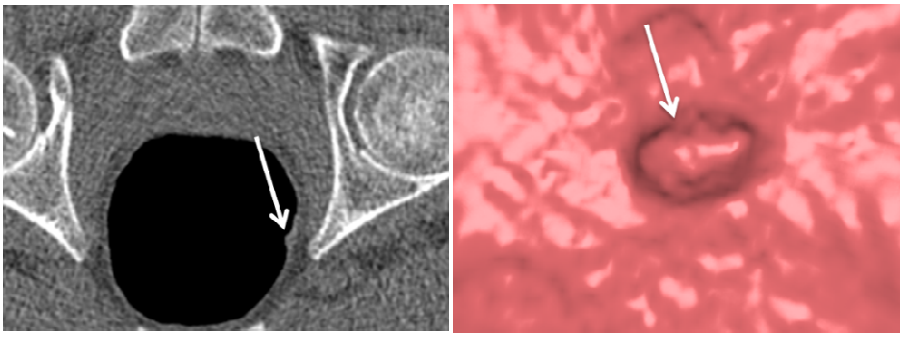


Fig. 1. Example of a lesion (9-mm adenoma, rectum) that was missed by both readers in the initial, unassisted reading, but was detected correctly with the use of CADe. (a) Axial view. (b) 3D endoluminal view.

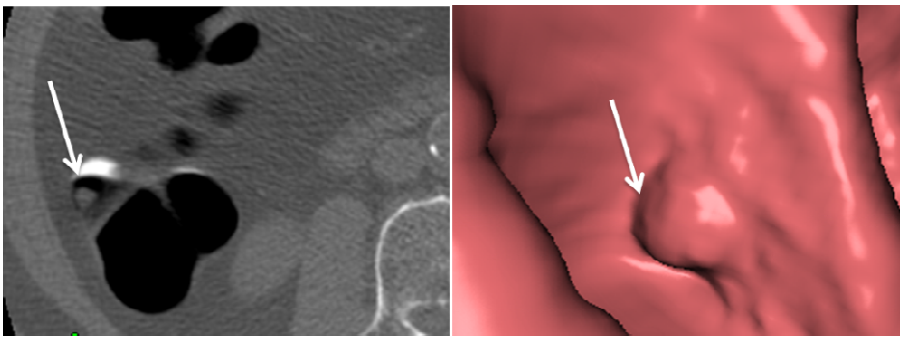


Fig. 2. An example of a lesion (7-mm adenoma, ascending colon) that was missed by both readers in the initial, unassisted reading, but was detected correctly with the use of CADe. (a) Axial view. (b) 3D endoluminal view.

For the flat lesions ≥10 mm in size, per-patient (per-lesion) sensitivities of the expert reader for CADe-unassisted and assisted readings were 59% [95% CI, 36-78%] (56%) and 71% [95% CI, 47-87%] (67%), respectively, whereas those of the non-expert reader were 41% [95% CI, 21-65%] (39%) and 47% [95% CI, 37-59%] (44%), respectively.

For 6-9 mm flat lesions, the corresponding per-patient (per-lesion) sensitivities of the expert reader were 59% [95% CI, 36-78%] (48%) and 76% [95% CI, 53-89%] (59%), respectively, whereas those of the non-expert were 47% [95% CI, 37-59%] (37%) and 82% [95% CI, 59-93%] (63%), respectively.

Fig. 1 and Fig. 2 show two examples of lesions that were missed by both readers in the initial (unassisted) readings, but detected correctly after reviewing the CADe prompts that correctly indicated the lesions. Fig. 3 shows an example of a significant lesion that was missed by both readers even though it was prompted correctly by the CADe system. In this case, however, both readers interpreted the lesion incorrectly as residual stool.

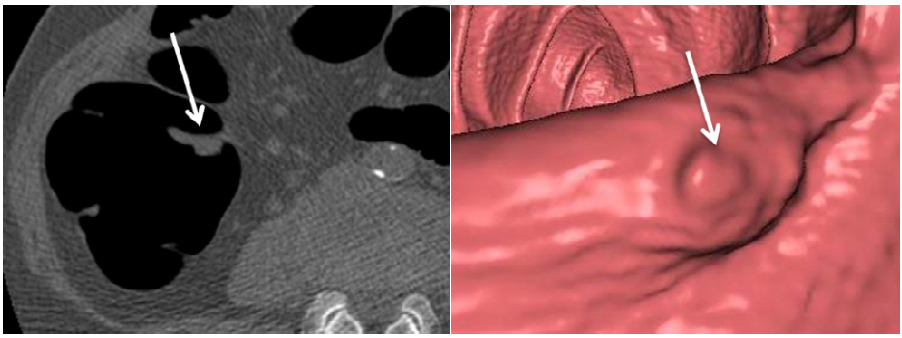


Fig. 3. Example of a lesion (14-mm adenocarcinoma, ascending colon) that was missed by both readers even though the CADe system prompted the lesion correctly to the readers. (a) Axial view. (b) 3D endoluminal view. Both readers considered this lesion as residual stool.

4 Discussion

This retrospective study demonstrated that the addition of CADe in CT colonography could substantially improve the identification of patients with non-polypoid flat lesions for both non-experienced and experienced readers. The per-patient sensitivity of the non-expert reader improved statistically significantly, while that of expert reader did not, indicating that CADe was more effective for non-expert than expert.

The readers did not always identify true lesions correctly even though they were marked correctly by CADe. If the readers had correctly identified all true flat lesions marked by CADe, they could have reached ≥90% detection sensitivity. Possible explanations for this suboptimal performance include the presence of untagged stool in some cases and possible lack of familiarity of the inexperienced reader to identify non-polypoid lesions.

A majority of the recognition errors was caused by the presence of untagged fecal residue. In this study, 2/3 of the patients had not received fecal-tagging agent. It should be noted that our study population was not optimized for latest state-of-the-art CT colonography but represented how CT colonography could be performed typically in daily clinical practice. In particular, the CT colonography protocols were probably not uniform across participating centers, and any motion artifacts or unsharpness of images could be detrimental for the detection of flat lesions.

There are several limitations in this study. First, we evaluate the effect of CADe only for a single expert and non-expert; thus, it is uncertain whether the results obtained in this study would generalize to a large pool of readers. Second, we did not include lesions that were detected on OC, but retrospectively invisible on CTC. Third, the study cohort was lesion-enriched and included a relatively small number of flat lesions.

The results of this study indicate that, although optimized CADe may be useful in improving the detection performance of radiologists for non-polypoid flat lesions, reliable detection of flat lesions also requires optimization of state-of-the-art CT colonography technique with fecal tagging, adequate distension, and high-resolution images.

5 Conclusion

The use of CADe optimized for the detection of flat lesions can substantially increase the sensitivity of human readers in the detection of flat lesions in a screening population.

Acknowledgments. This work was supported in part by grants of CA095279 and CA131718 from the NCI at the NIH.

References

1. ACS. Cancer Facts & Figures 2007. American Cancer Society (2007)
2. Winawer, S.J., Fletcher, R.H., Miller, L., et al.: Colorectal cancer screening: clinical guidelines and rationale. Gastroenterology 112, 594–642 (1997)

3. Macari, M., Bini, E.J.: CT colonography: where have we been and where are we going? Radiology 237, 819–833 (2005)
4. Johnson, C.D., Chen, M.H., Toledano, A.Y., Heiken, J.P., et al.: Accuracy of CT colonography for detection of large adenomas and cancers. N. Engl. J. Med. 359, 1207–1217 (2008)
5. Levin, B., Lieberman, D.A., McFarland, B., et al.: Screening and surveillance for the early detection of colorectal cancer and adenomatous Polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. CA. Cancer. J. Clin. 58, 130–160 (2008)
6. Soetikno, R.M., Kaltenbach, T., Rouse, R.V., et al.: Prevalence of nonpolypoid (flat and depressed) colorectal neoplasms in asymptomatic and symptomatic adults. JAMA 299, 1027–1035 (2008)
7. Park, S.H., Ha, H.K., Kim, A.Y., et al.: Flat polyps of the colon: detection with 16-MDCT colonography–preliminary results. Am. J. Roentgenol. 186, 1611–1617 (2006)
8. Pickhardt, P.J., Kim, D.H., Robbins, J.B.: Flat (nonpolypoid) colorectal lesions identified at CT colonography in a U.S. screening population. Acad. Radiol. 17, 784–790 (2010)
9. Fidler, J., Johnson, C.: Flat polyps of the colon: accuracy of detection by CT colonography and histologic significance. Abdom. Imaging 34, 157–171 (2008)
10. Yoshida, H., Dachman, A.H.: Computer-aided diagnosis for CT colonography. Semin. Ultrasound. CT MR 25, 419–431 (2004)
11. Yoshida, H., Dachman, A.H.: CAD techniques, challenges, and controversies in computed tomographic colonography. Abdom. Imaging 30, 26–41 (2005)
12. Yoshida, H., Nappi, J.: CAD in CT colonography without and with oral contrast agents: Progress and challenges. Comput. Med. Imaging Graph. 31, 267–284 (2007)
13. Regge, D., Della Monica, P., Galatola, G., et al.: Efficacy of Computer-aided Detection as a Second Reader for 6-9-mm Lesions at CT Colonography: Multicenter Prospective Trial. Radiology 266, 168–176 (2013)
14. Halligan, S., Mallett, S., Altman, D.G., et al.: Incremental benefit of computer-aided detection when used as a second and concurrent reader of CT colonographic data: multiobserver study. Radiology 258, 469–476 (2011)
15. Dachman, A.H., Obuchowski, N.A., Hoffmeister, J.W., et al.: Effect of computer-aided detection for CT colonography in a multireader, multicase trial. Radiology 256, 827–835 (2010)
16. Taylor, S.A., Iinuma, G., Saito, Y., Zhang, J., Halligan, S.: CT colonography: computer-aided detection of morphologically flat T1 colonic carcinoma. Eur. Radiol. 18, 1666–1673 (2008)
17. Näppi, J., Yoshida, H.: Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography. Acad. Radiol. 14, 287–300 (2007)
18. Regge, D., Laudi, C., Galatola, G., Monia, P.D., Bonelli, L., et al.: Diagnostic accuracy of computed tomographic colonography for the detection of advanced neoplasia in individuals at increased risk of colorectal cancer. JAMA 23, 2453–2461 (2009)

Application of Synthetic Sinogram Based Low-Dose CT Simulation and Fold-Preserving Electronic Cleansing Technique for CT Colonography

Chang-Won Kim[1], Jong-Hyo Kim[2,*], Hyunna Lee[3], Jeongjin Lee[4], Se-Hyung Kim[5], Zepa Yang[2], and Yeong-Gil Shin[3]

[1] Interdisciplinary Program of Bioengineering, Seoul National University College of Engineering, Seoul 151-742, Korea
kcw1110@snu.ac.kr
[2] Department of Intelligent Convergence Systems, Graduate School of Convergence Science and Technology, Seoul National University, Suwon-si, Gyeonggi-do 443-270, Korea
{yangzepa,kimjhyo}@snu.ac.kr
[3] The School of Computer Science and Engineering, Seoul National University, Seoul 151-742, Korea
hnlee@cglab.snu.ac.kr, yshin@snu.ac.kr
[4] The School of Computer Science and Engineering, Soongsil University, Seoul 156-743, Korea
leejeongjin@ssu.ac.kr
[5] Department of Radiology, Seoul National University College of Medicine, Seoul 110-744, Korea
shkim@radcom.snu.ac.kr

Abstract. This study evaluates the performance of the electronic cleansing technique in ultra-low-dose CT colonography, and explores the potential limit of the dose level which hampers the proper working of the electronic cleansing.

We applied a synthetic sinogram-based low-dose CT simulation technique to the DICOM images CT colonography studies to generate a set of ultra-low-dose CT colonography dataset ranging 20, 10, 5, and 1 mAs conditions. A fold-preserving electronic cleansing technique was applied to those ultra-low-dose CT colonography (CTC) dataset and the tagging cleansing performance was evaluated for each dose-level.

Keywords: ultra-low-dose CT colonography, low-dose CT simulation, electronic cleansing.

1 Introduction

While the use of computed radiography (CT) in modern imaging diagnosis is well established due to the advantages of 3-dimensional imaging of the human anatomy at relatively low cost and rapid imaging time, the increasing public concern to ionizing radiation is limiting its application in surveillance studies including CT colonography. Therefore, accomplishing the goal of minimizing the CT scan dose while maintaining

* Corresponding author.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 89–98, 2013.

diagnostic ability has been the subject of active studies in CT colonography. Recent studies reported that the detection performance of polyps could be maintained to an acceptable level even with ultra-low-dose (uLD) CT colonography examinations, which opened the possibility of using uLD scans in CT colonography for safer examinations in surveillance studies. However, we still need to know the margin of dose-level in uLD CT colonography where the polyp detection performance is maintained in order to ensure the diagnostic performance is not compromised. Also necessary is the verification of proper working of electronic cleansing (EC) technique in uLD scans which is expected to be routinely used in colon cancer screening with CTC.

This study evaluates the performance of the EC technique in uLD CTC with varying dose-levels in an attempt to find the limit of dose-level in which the proper working of the electronic cleansing technique is hampered. A low-dose CT simulation technique is applied to generate a set of uLD CTC data, which is then used to evaluate the performance of the newly developed EC technique.

2 Materials and Methods

2.1 Materials

Three CTC cases comprising prone and supine scans were downloaded from the CTC database in Seoul National University Hospital. The CTC cases were obtained using a 128-row CT scanner (Somatom Definition, Siemens, Erlangen) with scan conditions of 120 kVp, around 80 mAs, and 1mm slice thickness. The B30f kernel was used for image reconstruction.

A commercial CAD workstation (Xelis, Infinitt, Seoul, South Korea) was used for endoluminal visualization of colon. A newly developed electronic cleansing module was plugged into the CAD workstation and was evaluated in this study.

2.2 Methods

2.2.1 Overall Procedure

Shown in Fig. 1 is a schematic diagram of the overall process in this study. The regular-dose CTC images were input to the low-dose CT simulation program, which generated 4 sets of uLD CTC images for 20, 10, 5, and 1mAs levels. Each of regular-dose and uLD CTC images were loaded into the CTC CAD workstations, where the colon segmentation was carried out and colon mask was generated. The appropriateness of colon mask was first verified comparing segmented volume masks. The EC module was then launched to create the tagged image and cleansed image, which were evaluation in terms of appropriateness of tagging mask and image quality of endoluminal volume rendering. More detailed algorithm of low-dose simulation and EC are described in the following sub sections.

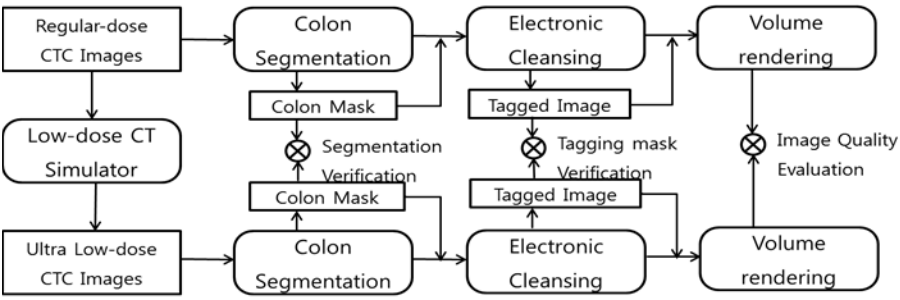


Fig. 1. Overall process of data processing and verification in this study

2.2.2 Simulation of Ultra-Low-Dose CT Imaging

The regular dose CT images (I_{high}) in DICOM format were used for the low-dose simulation. The Hounsfield unit (HU) of the images was converted to attenuation coefficients using the effective attenuation coefficient of water at each X-ray tube voltage. The attenuation coefficients were filtered to obtain the de-noised attenuation profile, and the attenuation profile ($A_s(g,d)$) for each gantry step was generated using the synthetic projection step which is a mathematical line integration of the attenuation coefficients from the point of the X-ray source to each detector element. The notation '*g*' means each gantry step in CT and '*d*' denotes the location in the detector array.

$$T_s(g,d) = e^{-A_s(g,d)} \tag{1}$$

The required geometry information for the synthetic projection step are the voxel size of the DICOM image, the distance between the source to the iso-center, the number of views, the reconstruction diameter, etc. Almost all the information is available in the DICOM header.

The attenuation profile was converted to transmittance by taking the exponential as in equation (2), followed by generation of the synthetic linear sinogram by multiplying the transmittance by the incidence flux (Q_0)and bow tie filter function ($b(d)$).

$$\overline{s_{s.high}}(g,d) = Q_0 \cdot B(d) \cdot T_s(g,d) \tag{2}$$

A synthesized linear sinogram was generated for each ‘g’ and ‘d’ by multiplying each standard deviation by the white Gaussian noise (WGN) and added to the synthetic linear sinogram. The white Gaussian noise has a zero mean and a standard deviation of one.

$$s_{sim}(g,d) = \overline{s_{s.high}}(g,d) + \sigma_{added}(g,d) \cdot WGN \tag{3}$$

The noise-added synthesized linear sinogram was then converted back to produce the simulated attenuation profile (A_{sim}), and was subtracted using the original attenuation profile, producing the added noise attenuation profile (A_{added}).

$$A_{sim}(g,d) = -\log\left(\frac{s_{sim}(g,d)}{Q_0 \cdot b(d)}\right) \tag{4}$$

$$A_{added}(g,d) = A_{sim}(g,d) - A_s(g,d) \tag{5}$$

This noise attenuation profile was filtered back projected and converted to HU to generate the added noise image (N_{added}). Finally, the noise image was added to the original high dose CT image (I_{high}), resulting in a simulated low dose CT image . The necessary parameters for the generation of additive quantum noise were obtain either from the literature or through experimental measurements.

2.2.3 Fold-Preserving Electronic Cleansing Method

This subsection describes the electronic cleansing method used in this study. A novel fold-preserving EC method was developed by using an adaptive fold reconstruction and tagged materials (TMs) separation technique [2]. The partial volume (PV) effect in CT produces unexpected soft-tissue-like (ST-like) layers between air and TM as well as aliasing artifacts between ST and TM after EC. Even though previous EC methods [4][5] resolved the PV effect, there still remains a common limitation: the degradation of ST structures caused by pseudo-enhancement (PEH) effect, which is an artificial increment of the observed CT density due to the presence of adjacent hyperdense contrast material. To address the partial volume (PV) and pseudo-enhancement (PEH) effects concurrently, material fractions and structural responses are integrated into a single reconstruction model. In our approach, colonic components including air, TM, interface layer between air and TM, and interface layer between soft-tissue (ST) and TM ($IL_{ST/TM}$) are first segmented. For each voxel in $IL_{ST/TM}$, the material fractions of ST and TM are derived using a two-material transition model [4]. And the structural response to identify the folds submerged in the TM is calculated based on the analysis of the Hessian matrix. Let the eigenvalues of Hessian matrix at a voxel be λ_1, λ_2, and λ_3 ($|\lambda_1| \le |\lambda_2| \le |\lambda_3|$). Based on this eigenvalue signature, structural response for submerged folds is defined using the rut-enhancement function [6] as follows:

$$F_{rut} = F_A \cdot F_B \tag{6}$$

with

$$F_A = \exp\left(-\frac{R_a^2}{2\alpha^2}\right), \quad R_a = \frac{|\lambda_1|}{\sqrt{\lambda_2 \lambda_3}} \tag{7}$$

$$F_B = \exp\left(-\frac{(R_b - \gamma)^2}{2\beta^2}\right), \quad R_b = \frac{|\lambda_2|}{|\lambda_3|}. \tag{8}$$

The parameters α and β control the enhancement range of folds and the parameter γ controls the sharpness of folds. In this paper, the parameters α, β, and γ are experimentally set as 0.7, 0.4, and 0.3, respectively. Then, the CT density value of each voxel in $IL_{ST/TM}$ is reconstructed based on both the material fractions and structural responses. The final CT density value I_{EC} is reconstructed by considering simultaneously the CT density value $I_{material-fractions}$ updated by the material fractions-based reconstruction model and the CT density value I acquired from the original CT images the weight factor of rut-enhancement function as follows:

$$I_{EC} = t_{ST} \cdot I_{ST} + t_{TM} \cdot \left\{\left(1 - \left(1 - F_{rut}\right)^{\kappa}\right) \cdot I_{TM} + \left(1 - F_{rut}\right)^{\kappa} \cdot I_{air}\right\}. \tag{9}$$

Let I_{ST} and I_{TM} represent the pure material densities of ST and TM, respectively. And let t_{ST} and t_{TM} $(=1 - t_{ST})$ represent the corresponding material fractions of each material. Let F_{rut} represent the value of rut-enhancement function at each voxel. The parameter κ controls the relative contribution of the rut-enhancement function to the fold preservation. In this paper, the parameter κ is experimentally set as 2. The material fractions remove the aliasing artifacts caused by PV effect in $IL_{ST/TM}$ effectively while the structural responses avoid the erroneous cleansing of the submerged folds caused by PEH effect.

3 Result

Fig. 2 shows example results of low-dose simulation technique. The three images of simulated uLD CT images for 20, 5, and 1mAs level exhibits the increasing level of noise and artifacts as the dose-level decreases. Especially, strong streak pattern arises from the high attenuating tagging material at the dose-levels of 5 and 1 mAs.

The colon segmentation results with the CAD workstation for the simulated uLD CTC dataset are shown in Fig. 3. It is clearly shown that the colon segmentation does not work properly due to the strong CT noise and artifacts with CAD workstation at dose levels below 5mAs.

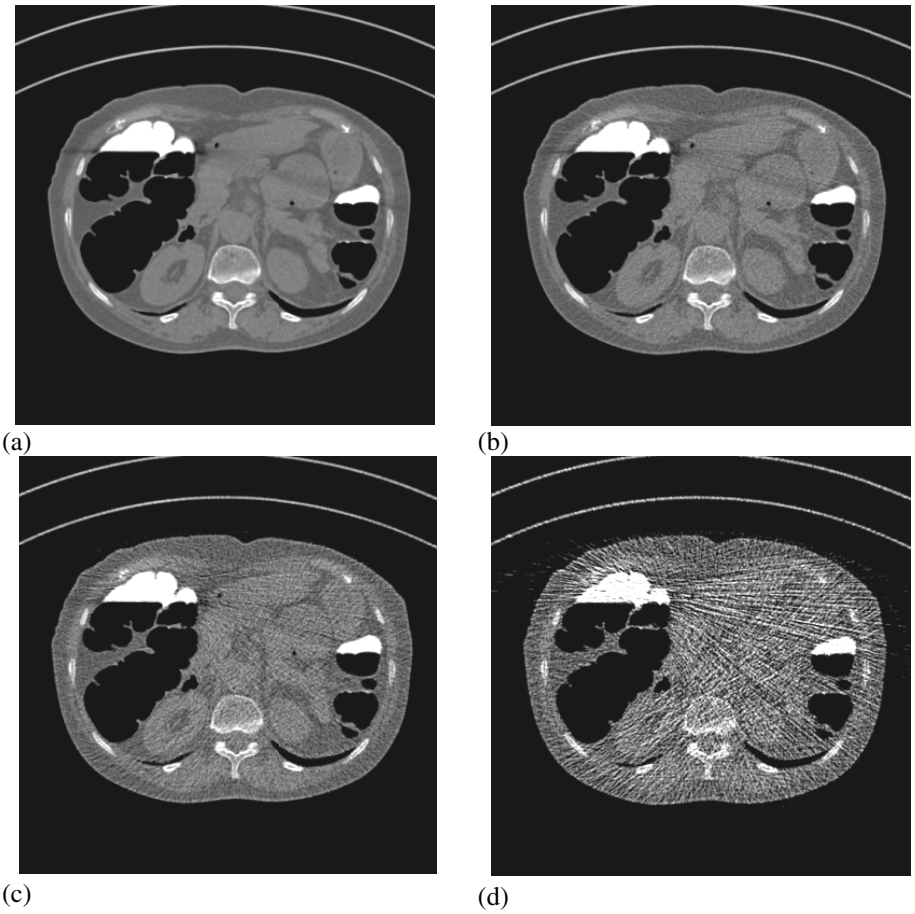


Fig. 2. Example results of the low-dose simulation technique applied to a CTC case with tagging material remaining in the colon. (a) real CT images of 80 mAs, (b)-(d) simulated uLD CT images for 20, 5, and 1 mAs levels.

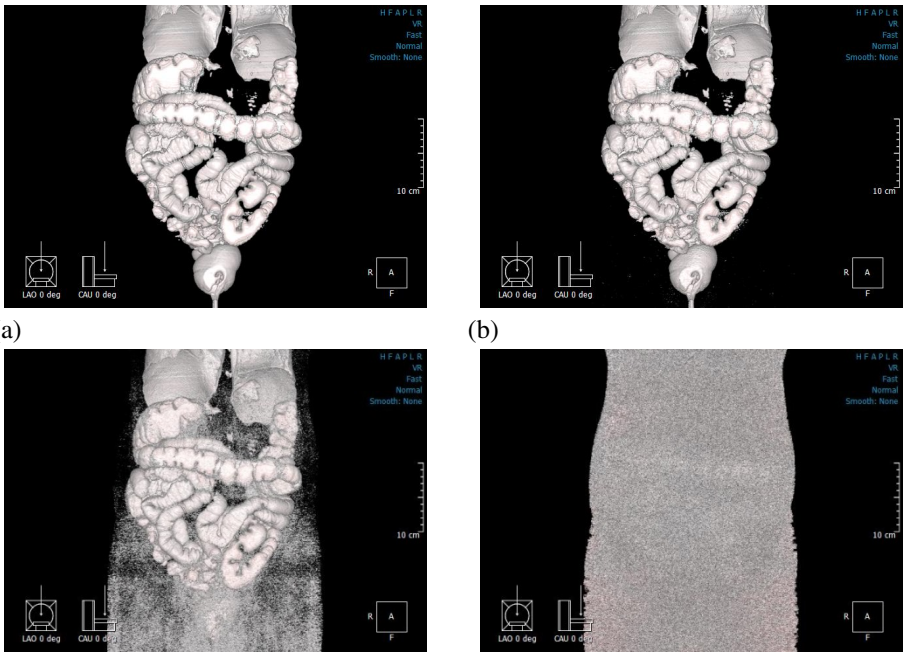


Fig. 3. Volume rendered view of the colon segmentation results with the CAD workstation. (a) real 80 mAs , (b)-(d) simulated uLD CT images for 20, 5, and 1 mAs,.

Examples of endoluminal volume rendering are shown in Fig. 4 for the uLD CTC images at varying dose-levels. Since the CAD workstation could not perform colon segmentation properly at 5 and 1 mAs levels, the colon paths were created with manual assist for those two dose-levels. At 20 and 10 mAs levels, the image quality was relatively acceptable although irregularities were observed on the folds and colon walls. However, at extreme low-dose levels structures of colon was no longer preserved due to excessive noise level.

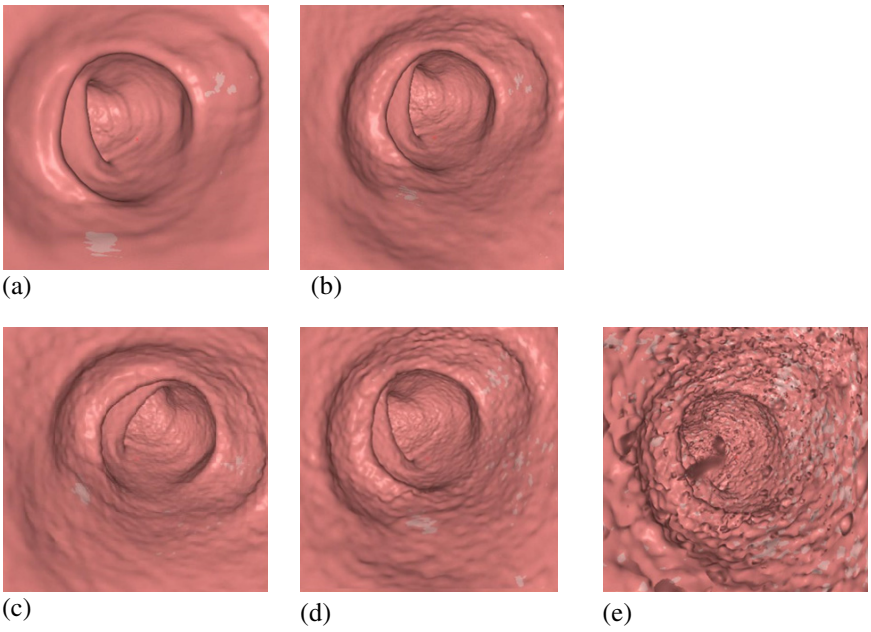


Fig. 4. Example endoluminal view of (a) original 80 mAs , (b)-(e) simulated 20, 10, 5, and 1 mAs images

Fig. 5 compares the endoluminal view of another CTC case at original 80mAs and simulated 10mAs level with and without EC. Although irregularities are noticed on colon walls and folds, the tagging material was well cleansed after applying EC at uLD CT images.

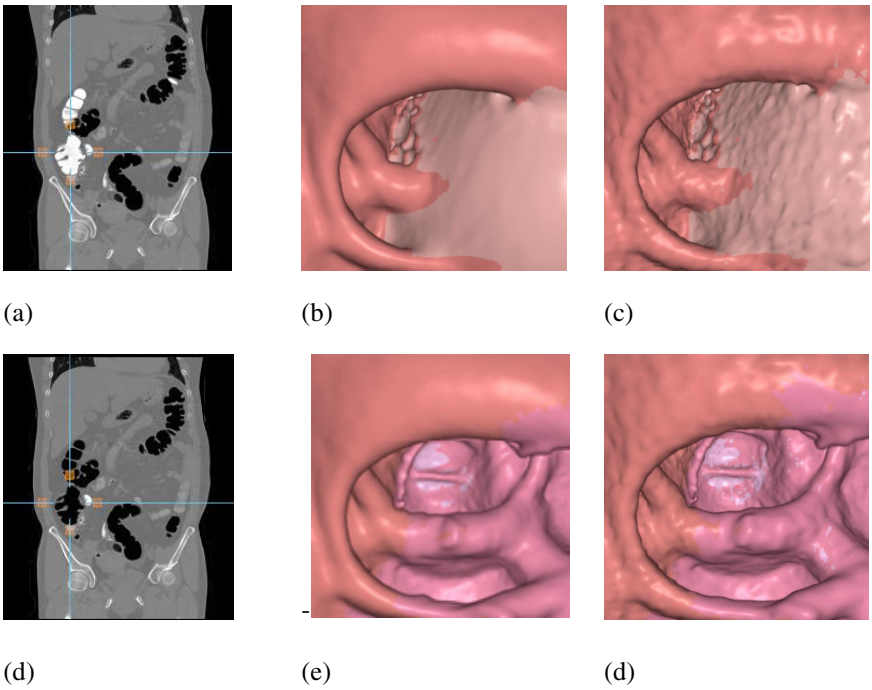


Fig. 5. (a) coronal view of original 80 mAs without EC, (b)-(c) endoluminal view of original 80 mAs and simulated 10 mAs without EC, (d) coronal view of original 80 mAs with EC, (e)-(f) endoluminal view of original 80 mAs and simulated 10 mAs with EC

4 Conclusion

This study evaluated the performance of EC technique in uLD scan conditions by applying a low-dose CT simulation technique. The CAD workstation could perform colon segmentation and electronic cleansing functions properly with uLD CT images down to 10mAs level. The performance of electronic cleansing and quality of endoluminal volume rendering was acceptable. However, the uLD CT images at extremely low dose levels at 5 and 1 mAs showed too strong noise and artifacts, and thus regarded not applicable to EC.

This study is limited in that only conventional FBP technique was used in low-dose simulation. Accordingly, there remains a possibility that EC could be performed with lower dose CT images obtained with advanced iterative reconstruction techniques.

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

References

1. Kim, C.W., Kim, J.H.: Application of CT Simulation Technique for Virtual Ultra-Low-Dose Trial in CT Colonography. In: Yoshida, H., Hawkes, D., Vannier, M.W. (eds.) Abdominal Imaging 2012. LNCS, vol. 7601, pp. 49–57. Springer, Heidelberg (2012)
2. Lee, H., Kim, B., Lee, J., Kim, S.H., Shin, Y.-G., Kim, T.-G.: Fold-preserving Electronic Cleansing using a Reconstruction Model Integrating Material Fractions and Structural Responses. IEEE Transactions on Biomedical Engineering 60(6), 1546–1555 (2013)
3. Wagner, R.F., Brown, D.G., Pastel, M.S.: Application of information Theory to the assessment of computer tomography. Med. Phys. 6(2), 83–94 (1979)
4. Serlie, I.W.O., Vos, F.M., Truyen, R., Post, F.H., Vliet, L.J.: Classifying CT image data into material fractions by a scale and rotation invariant edge model. IEEE Trans. Image Process. 16(12), 2891–2904 (2007)
5. Serlie, I.W.O., Vos, F.M., Truyen, R., Post, F.H., Stoker, J., Vliet, L.J.: Electronic cleansing for computed tomography (CT) colonography using a scale-invariant three-material model. IEEE Trans. Biomed. Eng. 57(6), 1306–1317 (2010)
6. 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(7), 3259–3277 (2008)

Iterative Reconstruction for Ultra-Low-Dose Laxative-Free CT Colonography

Synho Do, Janne J. Näppi, and Hiroyuki Yoshida

Massachusetts General Hospital Imaging and Harvard Medical School,
25 New Chardon Street, Suite 400C, Boston, Massachusetts 02114, USA
sdo@nmr.mgh.harvard.edu, jnappi@partners.org,
yoshida.hiro@mgh.harvard.edu

Abstract. Iterative reconstruction (IRT) makes it possible to acquire computed tomographic colonography (CTC) images at <50% of the radiation dose of conventional standard-dose filtered back-projection (FBP) without impairing image quality. It has also other significant advantages over FBP, including better noise suppression, reduction of image artifacts, and flexible mathematical formulation. In this pilot study, we explored the potential application of IRT in the implementation of an ultra-low-dose (ULD) laxative-free CTC examination. First, CTC images are reconstructed approximately with FBP to detect regions of fecal tagging and other high-density objects that can generate image artifacts. Next, the detected regions are projected to sinogram domain to guide the IRT process for the minimization of image noise, correction of beam-hardening artifacts, and virtual cleansing of fecal-tagged regions. For pilot evaluation, five patients were prepared for an ULD dual-energy CTC examination by use of non-cathartic dietary fecal tagging with iodine. For one patient, the CTC images were reconstructed by use of both the FBP and IRT methods. Preliminary results showed that the IRT-reconstructed images demonstrated superior image quality over the FBP-reconstructed images.

Keywords: Iterative reconstruction, non-cathartic, dose, radiation, virtual colonoscopy.

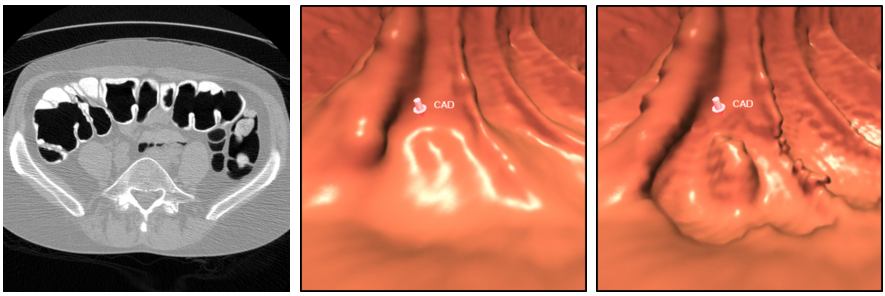
1 Introduction

The perceived inconvenience and discomfort of the cathartic bowel preparation of computed tomographic colonography (CTC) and colonoscopy have been identified as the primary causes of poor patient compliance in colon cancer screening [1]. Laxative-free CTC (lfCTC) has emerged as an alternative colorectal examination that does not use cathartic agents. A recent prospective multi-center clinical trial showed that lfCTC is effective in improving patient adherence by eliminating the discomforts of CTC examination [2].

A caveat of lfCTC is that it introduces large quantities of solid stool that adhere to colonic mucosa and present a multitude of shapes imitating those of

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 99–106, 2013.

colorectal lesions, thereby impairing the interpretation of lfCTC images. Conventional image-based electronic subtraction or electronic cleansing (EC) methods are not able to eliminate fecal-tagged stool precisely from CTC images, and their various subtraction artifacts that imitate small polyps and flat lesions tend to reduce detection performance especially on virtual endoscopic 3D images (Fig. 1) [2].



(a) (b) (c)

Fig. 1. (a) Laxative-free CTC data introduces large quantities of solid stool that adheres to colonic mucosa and presents a multitude of shapes. (b) Endoluminal 3D view of a 10-mm polyp covered by fecal tagging. (c) Conventional image-based electronic cleansing is able to reveal the polyp, but it also introduces subtraction artifacts that can distract the reader.

Another problem of CTC examinations is the perceived harmful radiation from CT scanners to patients [3]. However, iterative reconstruction (IRT) has recently made it possible to acquire CT images at less than 50% of the radiation dose of standard-dose CTC reconstructed with conventional filtered back-projection (FBP), without impairing image quality [4]. Together with automatic exposure control and low average current, the IRT may make it possible to perform ultra-low-dose (ULD) CTC at radiation doses of 1 mSv or less. This is substantially lower than the 4.4 – 7.6 mSv dose of conventional CTC [5] and the 2.4 mSv average yearly worldwide background radiation dose.

The IRT has also several other advantages over conventional FBP, including better noise suppression, reduction of image artifacts, and more flexible mathematical formulations that allow the inclusion of prior information. In particular, the IRT can be used to correct for beam-hardening artifacts that at low radiation doses appear in regions with metallic artifacts or with high concentrations of fecal tagging [6].

In this study, we investigated the application of IRT to ULD lfCTC. Patients were prepared for an lfCTC examination by use of dietary fecal tagging, and they were scanned at ULD with a dual-energy CT scanner. The CTC images were reconstructed with the IRT method and with the FBP method. The image qualities of the reconstructions were compared.

2 Methods

2.1 Iterative Reconstruction (IRT)

The conventional FBP method reconstructs CT images from raw sinogram-domain data in a single back-projection step, whereas the IRT reconstructs CT images iteratively by back-projection and forward projection between the sinogram domain and the image domain.

For IRT, the CT image reconstruction problem can be formulated as [7]

$$\hat{x} = \text{argmin}_x E_f(y, x) + \alpha E_r(x), \tag{1}$$

where $E_f(y, x) = ||y - Hx||^2$ is a data fidelity term via the projection between sinogram y and image x, $E_r(x) = ||\nabla x||^p$ is a regularization term (the prior), α is a weighting term, and H is the system matrix (the forward projector). To minimize the energy functional of Eq. (1), we use the lagged diffusivity fixed-point method, where the cost is approximated iteratively by a weighted quadratic cost followed by solving of the resulting linear normal equations using pre-conditioned conjugated gradient iterations [7]. The method is essentially a linearization technique for the nonlinear Euler-Lagrange equation that presents the total variation image restoration as a variational problem [8].

To implement the equations, the forward projector H and the back-projector are designed as a matched pair. The forward projection is based on a fast ray-tracing algorithm [9] that incorporates models for the flying focal spot, effective focal spot area, and the active detector area [10]. To ensure that the back-projector is matched to the forward projector, for each ray that is cast during the forward projection, we maintain a list of points of intersection and relative voxel weights between each ray and every voxel that it encounters. This ensures that the back-projection follows the exact same ray path in the reverse direction.

2.2 IRT-Based Correction for Image Noise

The noise in a CT scan can be modeled by a Poisson random process. The Poisson noise occurs due to the statistical error of low photon counts, thereby producing random thin bright and dark streaks that appear preferentially along the direction of greatest attenuation. With FBP, as the radiation dose is reduced, both the noise and image quality becomes worse. However, the statistical model of IRT reduces noise at each iteration by guiding the image reconstruction toward regularized images while preserving edges [7,6].

2.3 IRT-Based Beam Hardening Correction

Beam hardening produces dark streaks along the long axis of a single high-attenuation object or between two high-attenuation objects, such as iodinated contrast, barium, metal, and bone. As the x-ray passes through the body, low-energy x-ray photons are attenuated more easily than high-energy photons.

The dark streaks appear, because an imbalanced number of photons is measured in different views of the same region [6].

Modern CT scanners perform a beam hardening correction that assumes an average amount of beam hardening for the measured attenuation. However, materials with high atomic numbers, such as metal or iodine, cause a higher-than-average amount of beam hardening, and therefore their effect will not be fully corrected.

With IRT, the beam-hardening artifacts can be corrected. The first iteration reconstructs the CT images using uncorrected projection data. The regions that produced the beam-hardening artifacts have high radiodensity, and therefore they can be extracted by thresholding of the reconstructed CT images at a cutoff value of 200 Hounsfield units (HUs) [11]. Lower radiodensities do not cause beam hardening.

Next, the thresholded regions are projected forward into the sinogram domain to determine how much of the high-density material was present in each detector measurement. The projected data are then used to perform a custom beam-hardening correction for each detector element, where the corrected data are obtained by subtraction of the scaled error images from original reconstructed images.

2.4 IRT-Based Virtual Cleansing of Fecal-Tagged Regions

The IRT process can be used to perform virtual cleansing of the tagged materials on the reconstructed CTC images. In principle, the approximate regions of tagged materials can be identified by thresholding of high CT values as explained in Section 2.3. However, such thresholding excludes the artificial partial-volume interfaces that appear between tagged materials and air, because such interfaces have CT values similar to those of soft-tissue materials. Thus, such interface artifacts would remain in the final reconstructed CTC images.

To include the artificial partial-volume tagging interfaces together with the tagged materials in the IRT reconstruction process for the virtual cleansing, we use a gradient interface analysis method for detecting the interfaces [12]. Let T, S, and A denote the regions of tagging, soft tissue, and air, respectively, and let $T(x, v)$ denote the binary region that is obtained by thresholding of the image data x with a CT value of v. Furthermore, let parameters t_A and t_T denote the highest CT value of air ($t_A = -700$ HU) and lowest CT value of tagging ($t_T = 100$ HU), respectively. First, the CT values of x are clipped according to $v_1 = \min\{v, t_T\}$. Let x_1 denote the clipped image data. Next, the CT values of x are clipped according to

$$v_2 = \begin{cases} t_T + 700 & \text{if } v \geq t_T \\ v & \text{otherwise.} \end{cases} \tag{2}$$

Let x_2 denote the resulting data. Because the contrast between A and T in x_2, and that between S and T, is higher than the contrast between A and S, we can determine the artificial interface voxels between air and tagging according to

$$A|T = T(\nabla x_1, t_{\nabla x_1}) \cap T(\nabla x_2, t_{\nabla x_2}), \tag{3}$$

where $t_{\nabla x_1}$ is set to exceed the highest contrast of CT value within the interface $S|T$, and $t_{\nabla x_2}$ is set to exceed the highest contrast of CT values within the interfaces $A|S$ and $S|T$ [12].

In the IRT reconstruction process, the tagged regions and their partial-volume interfaces are projected to the sinogram domain and nullified by subtraction from the sinograms prior to the reconstruction of the CTC images [13]. This means that, instead of projecting rays through the tagged regions and their partial-volume interfaces, the rays are projected around the tagged regions. Therefore, the IRT-iterations based on Eq. (1) also perform implicit subtraction of the tagged materials and their partial-volume interfaces from the reconstructed CTC images.

3 Evaluation

Five patients were prepared for a laxative-free CTC examination by use of dietary fecal tagging with iodine. They were scanned in supine and prone positions at an effective radiation dose of 0.65 mSv with a collimation of 0.6 mm and tube currents of 40 mA (at 80 kVp) and 15 mA (at 140 kVp) by use of a dual-energy CT scanner (SOMATOM Definition Flash, Siemens Healthcare Global).

For one patient, the CTC images were reconstructed by use of both the conventional FBP method and the IRT method. The outcomes of the reconstructions were compared visually.

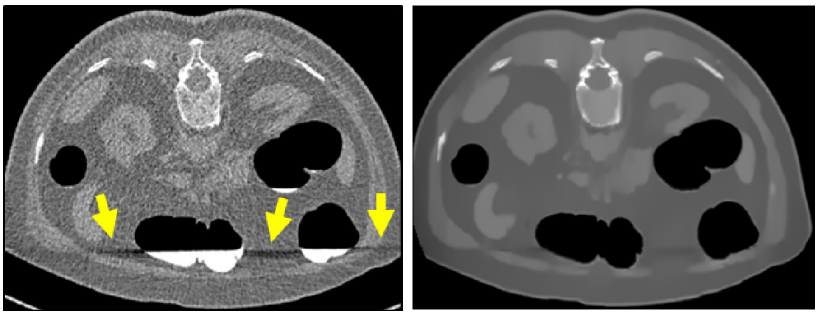
4 Results

Figure 2 shows a visual comparison of the image quality of conventional FBP-reconstructed and IRT-reconstructed axial CTC images. The FBP-reconstructed image presents a high level of image noise, and the regions of high-density tagging have produced severe beam hardening artifacts (arrows). In contrast, the IRT-reconstructed image shows significant suppression of the image noise and beam-hardening artifacts over the FBP. The image also demonstrates the implicit sinogram-based virtual cleansing of the fecal-tagged materials from the CTC image.

Figure 3 shows examples of the reconstructed images generated by the intermediate steps of the IRT reconstruction. The images show how the image noise and image artifacts are gradually minimized in the reconstructed images over the iterations.

5 Discussion

In this study, tagged regions were identified by use of approximate FBP-based image reconstruction. Another option would be to use the dual-energy information from the dual-energy CT scanner. However, most of the clinically available dual-energy CT scanners combine the information from both energy spectra in



(a) (b)

Fig. 2. Comparison of an (a) FBP-reconstructed and (b) IRT-reconstructed axial image from the same CTC scan. In (a), arrows indicate beam-hardening artifacts. In the IRT-reconstructed image, the image noise and beam-hardening artifacts have been suppressed with simultaneous virtual cleansing of tagged materials.

image domain to produce dual-energy-specific basis material images that are a linear combination of the images of each spectrum [14]. Such images are inferior to those obtainable by the sinogram-based approach, because the possibility to perform an exact material separation is lost. For exact material separation, sinogram-based methods must be applied.

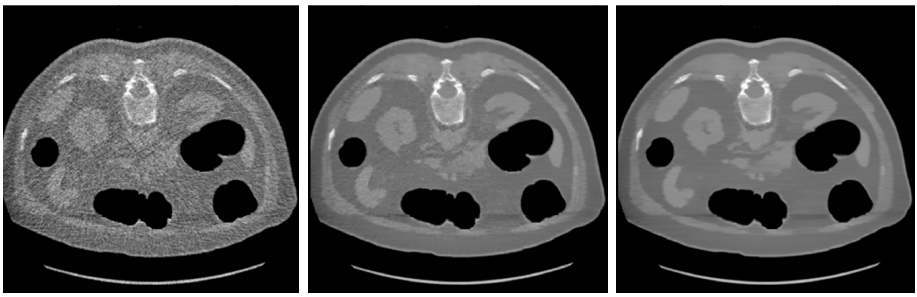


Fig. 3. From left to right: examples of the intermediate images calculated by the IRT

One of the limitations of conventional electronic cleansing techniques is that they can erroneously remove thin folds or parts of folds (Fig. 1). This can happen because beam hardening and pseudo-enhancement artifacts can artificially elevate the observed CT values of soft-tissue folds. However, because the IRT method involves beam-hardening correction of such artifacts, it has the potential advantage of preserving folds correctly in the reconstructed CTC images.

One of the limitations of the IRT method used in this study is that fecal-tagged regions are determined only once, prior to the reconstruction. A more practical approach would be to implement the segmentation step iteratively within Eq. (1):

during the first iterations, the fecal-tagged regions only need to be segmented approximately from low-resolution images, followed by progressive adjustment of the segmented region as the image resolution is gradually improved.

The IRT is a computationally demanding method, because it involves multiple back-projections and re-projections of large amounts of data. However, the computation time can be reduced substantially by parallelization of the calculations of ray projections by use of the non-recursive Siddon's algorithm [15].

Acknowledgements. This work was supported in part by grants of CA095279 from NCI at NIH. The authors also thank Daniele Regge and Simone Mazzetti (Institute for Cancer Research and Treatment, Candiolo, Torino, Italy) for providing clinical CTC data.

References

1. Gluecker, T., Johnson, C., Harmsen, W., Offord, K., Harris, A., et al.: Colorectal cancer screening with CT colonography, colonoscopy, and double-contrast barium enema examination: prospective assessment of patient perceptions and preferences. Radiology 227, 378–384 (2003)
2. Zalis, M., Blake, M., Cai, W., Hahn, P., et al.: Diagnostic accuracy of laxative-free computed tomographic colonography for detection of adenomatous polyps in asymptomatic adults: a prospective evaluation. Ann. Intern. Med. 156, 692–702 (2012)
3. Chang, K., Yee, J.: Dose reduction methods for CT colonography. Abdom. Imaging 38, 224–232 (2013)
4. Xu, Y., He, W., Chen, H., Hu, Z., Li, J., Zhang, T.: Impact of the adaptive statistical iterative reconstruction technique on image quality in ultra-low-dose CT. Clin. Radiol. (2013)
5. Boellaard, T., Venema, H., Streekstra, G., Stoker, J.: Effective radiation dose in CT colonography: is there a downward trend? Acad. Radiol. 19, 1127–1133 (2012)
6. Boas, F., Fleischmann, D.: CT artifacts: causes and reduction techniques. Imaging Med. 4, 229–240 (2012)
7. Do, S., Karl, W., Liang, Z., Kalra, M., Brady, T., Pien, H.: A decomposition-based CT reconstruction formulation for reducing blooming artifacts. Phys. Med. Biol. 56, 7109–7125 (2011)
8. Chan, T., Mulet, P.: On the convergence of the lagged diffusivity fixed point method in total variation image restoration. SIAM J. Numer. Anal. 36, 354–367 (1999)
9. Siddon, R.: Fast calculation of the extract radiological path for a three-dimensional CT array. Med. Phys. 12, 252 (1985)
10. Do, S., Cho, S., Karl, W., Kalra, M., Brady, T., Pien, H.: Accurate model-based high resolution cardiac image reconstruction in dual source CT. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 330–333. IEEE Press, Boston (2009)
11. Hsieh, J., Molthen, R., Dawson, C., Johnson, R.: An iterative approach to the beam hardening correction in cone beam CT. Med. Phys. 27, 23–29 (2000)

12. Näppi, J., Yoshida, H.: Fully automated three-dimensional detection of polyps in fecal-tagging CT colonography. Acad. Radiol. 25, 287–300 (2007)
13. Do, S., Liang, Z., Karl, W., Brady, T., Pien, H.: A projection-driven pre-correction technique for iterative reconstruction of helical cone-beam cardiac CT images. In: Proc. SPIE Medical Imaging: Physics of Medical Imaging, vol. 6913. SPIE Press (2008)
14. Brooks, R.: A quantitative theory of the Hounsfield unit and its application to dual energy scanning. J. Comput. Assist. Tomogr. 1, 487–493 (1977)
15. Jacobs, F., Sundermann, E., De Sutter, B., Christiaens, M., Lemahieu, I.: A fast algorithm to calculate the extract radiological path through a pixel or voxel space. J. Comput. Inf. T. 6, 89–94 (1998)

Global Colon Geometric Structure Analysis Based on Geodesics and Conformal Flattening

Hao Peng[1], Rui Shi[1], Shiqing Xin[2], and Xianfeng Gu[1]

[1] Department of Computer Science, Stony Brook University, Stony Brook, NY 11794-4400
[2] Institute of Computer Science & Technology, Ningbo University, China, 315211
hao.peng@stonybrook.edu, {rshi,gu}@cs.sunysb.edu,
xinshiqing@gmail.com

Abstract. Global geometric structure analysis plays an important role in computer vision and medical imaging. Human colon has complex geometric structures; colonic modeling has been the most challenging issue for computer aided detection and diagnosis (CADe and CADx). The increasing demand for colon modeling relies on the construction of an accurate endoscopy view for physician to locate polyps, precursors of colorectal cancer. This work focus on automatically locating the Teniae Coli, Haustral Folds and extracting centerline, which gives the global geometric structure of a colon wall surface anatomy. A series of algorithms based on geodesics and conformal flattening are proposed: *auxiliary Riemannian metric* algorithm for Teniae Coli tracking; *geodesic clustering* method for Haustral folding location; *harmonic mass center* method for centerline construction. Our method is fully automatic, accurate and robust. We tested our method on real colon surfaces reconstructed from CT images. The experimental results demonstrate the efficiency and efficacy of our method.

Keywords: Colon Geometric Structure, Colonic Modeling, Geodesics, Conformal Flattening, Colon Centerline, Teniae Coli, Haustral Folding.

1 Introduction

Colorectal carcinoma is the third leading cause of cancer death and the third most commonly diagnosed cancer [1]. Since it is highly preventable, it is extremely important to detect and treat colorectal cancer in the earliest stage. Computed Tomographic (CT) based Virtual Colonoscopy has been emerging as a reliable, non-invasive technique for colon cancer screening [12]. This technique reconstructs a three-dimensional colon structure and virtual endoscopic layout within the colonic lumen from CT volume data.

Human colon has complex structures because colon often turns, twists and even mobiles in various directions. Therefore, colonic modeling has been the most challenging issue for computer aided detection and diagnosis (CADe and CADx). The ultimate goal of colon modeling is to construct an accurate endoscopy view for physicians to locate polyps, which has been discovered as early signs of colorectal cancer. Thus, accurate extraction of Teniae Coli muscle and centerline is in emerging need. Teniae Coli (TC) are three separate longitudinal ribbons of smooth muscle on the surface of colon. They are parallel, equally distanced and form a piece wise triple helix structure from ascending to sigmoid colon. TC muscle, with tension, contracts lengthwise to produce haustra

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 107–116, 2013.

folds, the bulges in the colon. Furthermore, haustral folds represent folds of mucosa within the colon. They are formed by circumferential contraction of the inner muscular layer of the colon.

This work focus on automatically locating the Teniae Coli and Haustral Folds and extract centerline, which gives the global geometric structure of a colon wall surface anatomy.

In order to solve the problem, we apply two geometric techniques: surface geodesics and conformal flattening. From colon anatomy, the three muscle ribbons of Teniae Coli are stiffer, therefore regions near Teniae Coli are relatively flatter. We can locate the Teniae coli by tracing longitudinal geodesics. Hasutral Folds traverse the Teniae Colis, and are with negative Gaussian curvatures, each Haustral fold contains a geodesic loop. Therefore, we locate Haustral folds by searching latitudinal geodesic loops. The Centerline goes through the centers of geodesic loops of Haustral folds, and is parallel to the Teniae Colis.

This method encounter difficulties in the intersection areas between the ascending and transverse colons, and the transverse and descending colons, where the longitudinal geodesics along TCs are very close to one another, small geometric noises may make them cross each other and produce mismatch. We tackle this difficulty by using conformal colon flattening technique. We flatten the colon surface onto a planar rectangle, and separate them on the parameter domain. Anatomically, the TCs should be parallel to one another and never cross. Although they are close to one another in $\mathbb{R}^3$, they remain far apart on the parameter domain.

Contributions. This work proposes a novel method for colon wall surface global geometric structure analysis based on geodesics and colon flattening. We design an *auxiliary Riemannian metric* algorithm for Teniae Coli tracking;

The method is capable of finding Teniae Coli, Hausstral Folds, and extracting centerline. The method is fully automatic, global and robust to geometric noises.

2 Previous Works

Colon Flattening. Various colon flattening methods [3][4][15] have been proposed, among which conformal mapping algorithm showed the advantages of preserving local feature and minimizing the distortion by maintaining all the angles. Conformal Mapping has successfully been presented on colon structure, such as Colon Wall Flattening [4], Fold Segmentation and Supine-Prone colon registration [18].

Geodesics. Geodesic distance has always been a powerful tool for surface modeling. However, only a few researchers have applied this method in the field of medical imaging in the past due to the complexity of the algorithm. Sharir and Schorr [9], Mitchell et al.[10], Chen and Han [2], Xin and Wang [16] have proposed different algorithms to solve discrete geodesic problem, while the complexity of the algorithms has been lowered a lot, which gives us an accurate and feasible measuring feature for colonic modeling.

Colon Modeling. Previous work have devoted to manually [5], semi-automatically [7][6] or fully automatically [20] extracting the three TCs. However, most of the methods are mainly focusing on local shape analyzing, for instance, curvature based TC extraction. Centerlines are mainly used for simulating the traverse route of camera in Virtual Colonoscopy. Topological thinning [17] technique is traditionally assumed to provide the most accurate centerline result by peeling off a volumetric object layer iteratively to get the central layer. The repetitive procedure of layer-by-layer removal is time consuming while preserving the topology of the object. A faster centerline extraction [14] has been presented based on distance field in the colon lumen. Haustrals and Haustral folds are major substructure forming the colon inner wall. Zhu[19] brought forward a robust segmentation method and Hampshire et al[13] have proposed a surface registration method based on haustral fold matching.

3 Algorithm

In this section we present our algorithms in details, include the surface geodesic, conformal mapping and Teniae Coli/Centerline extraction algorithms.

Discrete Geodesic on Surface. In this paper, we use an improved version of Chen-Han's algorithm, which was developed in [16] and has been well know for it's good performance.

Let S be a (triangulated) polyhedral surface in R^3, defined by a set of faces, edges and vertices. Without loss of generality, we assume that S has a complexity of n since Euler's formula affirms that the number of vertices, edges and faces of a polyhedral surface are linearly related [11]. Suppose P is a path restricted on the surface S and this path goes through a sequence of vertices and edges. We call this sequence Γ.

In studying shortest path problem, we need to define a well-designed data structure called a window. A window encodes a set of shortest paths that goes through the same sequence of vertices and faces [2]. Since a typical edge vertex sequence Γ ends with a vertex or an edge, we classify windows into two types: (1) pseudo-source windows and (2) interval windows; see Fig. 1. Technically, we use a pair (d, v) to denote the key information encoded in a pseudo-source window at vertex v, where d is the shortest distance from the source to v restricted on the sequence. If Γ ends with an edge e, we call the last vertex r in Γ a root. Sharir and Schorr [9] proved that the point set $\{p|p \in e$, the shortest path from r to p can be unfolded into a straight line segment$\}$ is an interval (see Fig. 1). So we use a four-tuple $(d, l_r, e, [a, b])$ to denote an interval window, where r is the root, l_r is the unfolded image of vertex r, and d is the shortest distance from the source to r restricted on the edge-vertex sequence of interest.

To avoid the combinatorial explosion of the number of windows, Xin and Wang [16] used two important observations, namely "one angle one split" and "Filtering theorem".

One Angle One Split. As shown in Figure 3(a), w_1 and w_2 are two interval windows that cover the same angle $\angle v_1 v v_2$. According to the window derivation algorithm in [2], both w_1 and w_2 may have two children, which lead to four new windows $2w_1^1$, w_1^2, w_2^1, w_2^2. However, Chen and Han one angle one split theorem [2] states that at least one of the four children does not help when determining shortest paths. That is to say, among all the windows covering the same angle, at most one of them can have

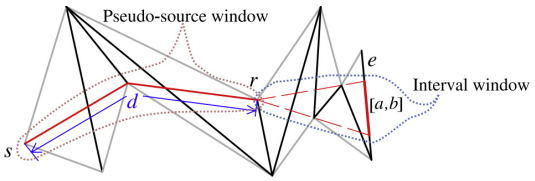


Fig. 1. A pseudo-source window encodes a shortest path from the source point to a vertex, while an interval window encodes a set of shortest paths that share the common edge sequence

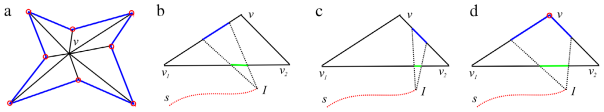


Fig. 2. Computing the children of a window: (a) A pseudo-source window at a saddle vertex v may have children on opposite edges and at adjacent vertices; (b) An interval window w on edge v_1v_2 may have one or two interval-window children, depending on how the line segment l_v intersects w's interval

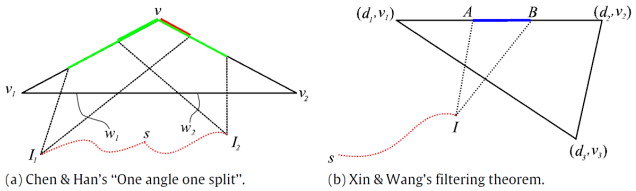


Fig. 3. Two key observations :(a) Chen and Han observed that only three of the four possible edge sequences in this example can provide shortest paths. The three green windows belong to edge sequences that will continue, and the red one corresponds to an edge sequence that must stop because the shortest paths that it contains cannot continue to pass this triangle. (b) Xin and Wang [16] suggested filtering out useless windows using the current estimates of the distances to the vertices.

two children that possibly determine a shortest path. Thus, one associates each angle with a winning window, by which one can check the validity of new windows. The CH algorithm guarantees that the window tree is of $O(n^2)$ in size and $O(n)$ in depth.

Filtering Theorem. Xin and Wang [16] observed that the majority of windows created by the CH algorithm [2] are useless. Therefore they proposed a filtering theorem to remove those useless windows. As shown in Figure 3(b), w is an interval window on the edge v_1v_2 and the unfolded image of w's root is located on the plane of $\triangle v_1v_2v_3$.

Discrete Conformal Flattening The colon wall surface is a topological cylinder, denoted as S, with two boundary loops $\partial S = \gamma_0 - \gamma_1$.

First, we compute a harmonic function $f : S \to \mathbb{R}$, with Dirichlet boundary condition, this can be achieved by the Laplace equation:

$$\begin{cases} \Delta f & = 0 \\ f|_{\gamma_0} & = 0 \\ f|_{\gamma_1} & = 1 \end{cases}$$

Fig. 4. Holomophic 1-form colon flattening. From left to right, the holomorphic 1-form, the closed harmonic 1-form ω, the exact harmonic 1-form df, the conformal flattening image.

In discrete case, the function f is defined on vertices, the discrete Laplace-Beltrami operator is defined as follows, assume $v_i \notin \partial S$ is an interior vertex,

$$\Delta f(v_i) = \sum_{[v_i,v_j]\in S} w_{ij}(f(v_j) - f(v_i)),$$

where w_{ij} is the edge weight for edge $[v_i, v_j]$,

$$w_{ij} = \frac{1}{2}(\cot\alpha + \cot\beta),$$

where α and β are the two corner angles against the edge on the mesh. The df is a discrete harmonic 1-form, $df([v_i, v_j]) = f(v_j) - f(v_i)$.

Second, we compute the generator of the first cohomology group $H^1(S, \mathbb{R})$. The surface is a cylinder, therefore, H^1 is of one dimensional. Find a shortest path connecting the two boundary loops γ. Slice the surface along γ to get a topological rectangle $\bar{S}$, with boundary segments $\gamma^+, \gamma_0, \gamma^-, \gamma_1$. Define a function $g : \bar{S} \to \mathbb{R}$,

$$g(v_i) = \begin{cases} 1 & v_i \in \gamma^+ \\ 0 & v_i \in \gamma^- \\ rand\ otherwise \end{cases}$$

Then along γ^+ and γ^-, dg equals to 0. Hence $\eta = dg$ is a well defined closed 1-form on the original mesh S. η is the generator of $H^1(S, \mathbb{R})$.

Third, we compute the unique harmonic 1-form ω cohomological to η, $\omega = \eta + dh$, where h is a 0-form. This requires to solve the following Possion equation,

$$\delta\omega = \delta(\eta + dh) = 0.$$

In discrete case, this is equivalent to solve a linear equation system,

$$\sum_{[v_i,v_j]\in S} w_{ij}(\eta[v_i, v_j] + h(v_j) - h(v_i)) = 0, \forall v_i \in S.$$

Both df and ω are harmonic 1-forms. It can be shown that

$$^*\omega = cdf,$$

where c is the inverse of the harmonic energy of f, given by

$$c^{-1} = \int_S \|\nabla f\|^2 = \sum_{[v_i,v_j]\in S} w_{ij}(f(v_j) - f(v_i))^2.$$

therefore we get a holomorphic 1-form $\omega + \sqrt{-1}^*\omega$.

Finally, the conformal flattening is given by the integration of the holomorphic 1-form on the sliced mesh $\bar{S}$. Choose a base vertex $v_0 \in \bar{S}$, for any vertex $v_i \in \bar{S}$, choose a path in $\bar{S}$ consisting of a sequence of oriented edges $\{e_0, e_1, \cdots, e_k\}$, then

$$\phi(v_i) = \int_{v_0}^{v_i} \omega + \sqrt{-1}^*\omega = \sum_{j=0}^{k} \omega(e_j) + \sqrt{-1}^*\omega(e_j).$$

Auxillary Riemannian metric algorithm for Teniae Coli Extraction The Teniae Coli extraction is based on geodesic tracking method. Basically, given two boundary points $p \in \gamma_0$ and $q \in \gamma_1$ of the colon surface, we compute the geodesic connecting them. In practice, the geodesic under the original induced Euclidean metric may become spirals, and circle around the colon surface for several cycles. In order to "straighten" the geodesics, we design a special Riemannian metric to penalize the swirling.

The conformal mapping $\phi : S \to \mathbb{R}^2$ maps the colon surface onto the (u, v) plane. The iso-u curves on the surface are the longitudes, the iso-v curves are the latitudes. The surface has the original induced Euclidean metric $\mathbf{g}$, because the mapping is conformal, therefore

$$\mathbf{g} = e^{2\lambda(u,v)}(du^2 + dv^2).$$

Also, we would like the geodesic to align with the longitude, then we define a new Riemannian metric as

$$\mathbf{g} = e^{2\lambda(u,v)}(du^2 + \lambda dv^2), \tag{1}$$

here the parameter $\lambda > 1$ is greater than 1, to penalize any movement along the latitudes. In practice, we found λ value around 1.5 gives correct results.

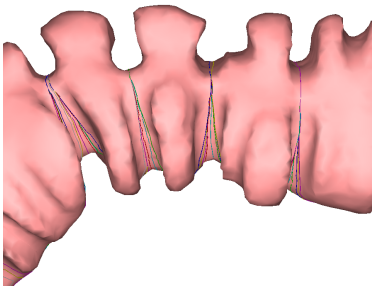


Fig. 5. Geodesic clustering for locating Haustral folds. Geodesic loops (colored curves) concentrate on the Haustral fold regions, each cluster corresponds to one fold.

Geodesic Clustering algorithm for Haustral Folding Location. The Haustral folding can be located by latitudinal geodesic loops. First, the colon surface S is sliced along a path γ connecting two boundary loops γ_0 and γ_1, the sliced surface $\bar{S}$ is conformally mapped to a rectangle on the plane by ϕ, with four boundary edges γ^+, γ_0, γ^- and γ_1, $\phi(\gamma^+)$ is $u = 0$, $\phi(\gamma^-)$ is $u = 1$.

We uniformly sample γ^+, the samples are $\{p_i^+\}, i = 1, 2, \cdots, n$. Each sample point p_i^+ has a unique corresponding point p_i^- on γ^-, such that p_i^+ and p_i^- coincide on the original colon surface S.

Through each pair of (p_i^+, p_i^-), compute a geodesic path on $\bar{S}$ using the original colon surface Riemannian metric, which is a geodesic loop τ_i. The geodesic loops tends to converge in the Haustral folds as shown in Figure 5.

Then we draw a vertical line Γ, $u = 0.5$. Suppose Γ intersects the geodesic loop τ_i at q_i. By clustering the intersection points q_i's on the line Γ, we cluster the geodesic loops τ_i. Each cluster of geodesics corresponds to a Haustral fold.

For each cluster on Γ, we pick the center point, denoted as $\{Q_1, Q_2, \cdots, Q_k\}$. Finally, we slice the original colon surface along Γ, each center point Q_k is split to two dual points Q_k^+ and Q_k^-. We compute the geodesic connecting Q_k^+ and Q_k^-, this gives us a geodesic loop, which indicates the corresponding Haustral fold.

Harmonic Mass Center Algorithm for Centerline Extraction. From the above algorithms, each Haustral Fold contains a geodesic loop. In this step, we compute the *harmonic mass center* for each geodesic loop, and connect these centers to construct the Centerline.

Suppose a Haustral fold geodesic loop is represented as a sequence points in $\mathbb{R}^3$, $\{p_1, p_2, \cdots, p_n\}$, we use principle component analysis method to find the best fitting plane π. Then we project all the points to the plane, the projection images are denoted as $\{\bar{p}_1, \bar{p}_2, \cdots, \bar{p}_n\}$, which form a planar polygon P.

It is possible that the polygon P is concave, therefore, simple mass center of P may be outside the colon wall surface. Instead, we invented the concept of *harmonic mass center*, which is guaranteed to be inside the colon surface.

Let Ω be a planar Jordan domain, there is a unique harmonic map $\phi : \Omega \to \mathbb{D}$, with Dirichlet boundary condition,

$$\begin{cases} \Delta\phi &= 0 \\ \phi|_{\partial\Omega} &= e^{i\theta(s)}, \end{cases}$$

where $\theta(s) = \frac{2\pi s}{C}$, where s is the arc length parameter of the boundary curve $\partial\Omega$, C is the total length. The computational algorithm is very similar to that of computing the harmonic functions on the colon surface.

The *harmonic mass center* of Ω is given by $\phi^{-1}(0)$. By connecting the harmonic mass centers of all consecutive Haustral fold geodesic loops, the Centerline of the whole colon surface can be constructed.

4 Experimental Results

Data Preparation The proposed method was applied to a CTC database including 50 patient studies collected with IRB approval. The patients were aged from 50 to 80 years.

Each patient was scanned at both supine and prone positions by 4- and 8-MDCT scanners (Light Speed Ultra, GE Medical Systems, Milwaukee, WI), resulting in 100 CT scans. The scanning protocol included mAs modulation in the range of 120-216 mA with kVp of 120-140 values, 1.25-2.5 mm collimation and reconstruction interval of 1 mm. The slice thickness of the CTC images ranged from 0.96 to 1.25 mm, and the in-plane pixel size from 0.53 to 0.76 mm. In the database, a total of 84 clinically significant polyps and masses (larger than 30 mm), sized in the range of 6-35 mm, were confirmed by both optical and virtual colonoscopies. The CT images are segmented using MAP(maximum a posteriori) method, and reconstructed using marching cube method. The reconstructed meshes have many sprue handles, by topological denoising, all the fake handles are removed. Related work demonstrated by Liang[9] and Hong[4].

Colon Flattening. The colon meshes have approximately 100k triangular faces. The flattening process mainly involves solving Laplace equations, which is converted to solve large sparse linear systems. We use preconditioned conjugate gradient method to solve them based on Eigen numerical package. It takes about 2.5 minutes for flattening one colon surface. In order to verify the mapping is bijective, we compute the signed area of all triangles on the parameter plane. In all testing cases, there is no flipped faces. Fig. 4 demonstrates the colon flattening process. In order to visualize the conformality, we put checker a board texture on the parameter domain, then the texture mapping result shows the mapping preserves all the right corner angles of checkers.

Teniae Coli Extraction. Teniae Coli extraction is carried out by tracing longitudinal geodesics. First, we conformally map the colon surface to a planar rectangle, then uniformly sample both the top and bottom boundary loops on the parameter plane, $\{p_1, p_2, \cdots, p_k\}$ and $\{q_1, q_2, \cdots, q_k\}$, where p_i and q_i are on the same longitude. For each pair $\{p_i, q_i\}$, we trace a geodesic using the Auxillary Riemmannian metric in Eqn.1, where $\lambda = 1.5$. For each colon surface, the Teniae Coli extraction time is about 30 sec. Figure 6 shows the Teniae Coli extracted for a whole colon surface.

Haustral Fold Location. First, we uniformly sample the left vertical boundary edge and obtain about 300 points, find their correspondences on the right boundary edge. Through each pair of sample points, we compute a geodesic loop. The geodesic loops intersects the longitude in the middle. Then we cluster the loops using k-mean clustering method on these intersection points, where we choose k to be about 30. Then through the mass center of each cluster of intersection points, we compute a geodesic loop, which indicates a Haustral Fold. This process takes about 50 seconds. Figure 5 shows the geodesic clustering on a local region of the colon. Different geodesic loops are encoded by different colors. The cut of the colon is at the bottom, all loops converge at the top.

Harmonic Mass Center for Centerline Construction. We run PCA to compute the best fitting plane for each geodesic loop, and project the loop to the fitting plane to get a planar polygon. Then we use Ruppert's Delaunay refinement method to tessellate the planar polygon, using harmonic map to transform it to the planar disk. The preimage of the disk center is the harmonic mass center. By connecting the harmonic mass centers, we obtain the centerline. This process takes about 1.5 minutes.

Validation. We validate our computational results by comparing with those obtained by human experts. The computed Teniae Coli, Hautral fold and centerlines are very close to those labeled by the experts, and are satisfactory for practical applications in virtual colonoscopy. However, no exact metrics have been applied for comparison.

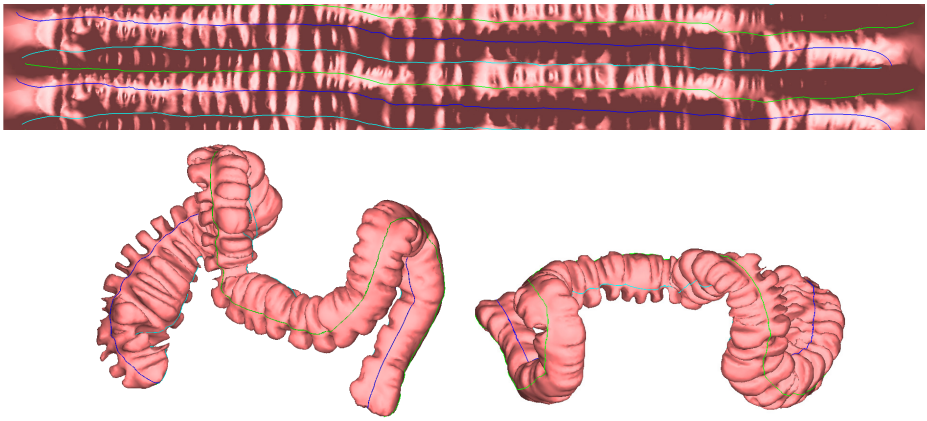


Fig. 6. Teniae Coli tracking results on the colon surface with different views (bottom) and the flattened image (top). The flattened view show two periods to visualize the whole Teniae Coli structures.

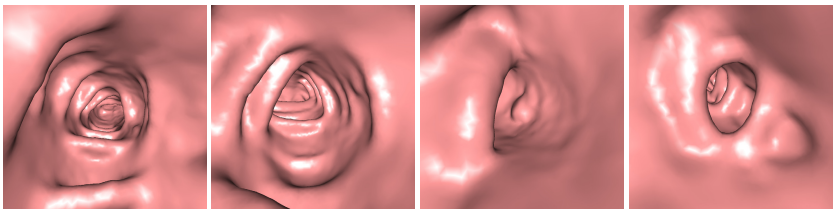


Fig. 7. Virtual colonosocpy guided by the extracted center line and polyps located

5 Conclusion and Discussion

This work focus on automatically locating the Teniae Coli, Haustral Folds and extracting centerline, which gives the global geometric structure of a colon wall surface anatomy. The method is based on surface geodesics and conformal flattening. Auxiliary Riemannian metric method is proposed for guiding global geodesics to trace Teniae Coli; Geodesic clustering algorithm is designed for Haustral folding location; Harmonic mass center method is applied for centerline construction. The algorithm pipeline has been tested on colon scans from real patients. It is automatic, robust and accurate. Current method carries out all the computation only on one colon surface. In practice, each patient was scanned at both supine and prone positions. We will explore the approach to combine two colon surfaces to improve the robustness and accuracy.

References

1. The American Cancer Society: Colorectal cancer facts & figures 2011-2013 (2011)
2. Chen, J., Han, Y.: Shortest paths on a polyhedron. In: Proceedings of the Sixth Annual Symposium on Computational Geometry, SCG 1990, pp. 360–369 (1990)
3. Haker, S., Angenent, S., Tannenbaum, A., Kikinis, R.: Nondistorting flattening maps and the 3d visualization of colon ct images. IEEE Transactions on Medical Imaging 19, 665–670 (2000)
4. 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)
5. Huang, A., Roy, D., Franaszek, M., Summers, R.M.: Teniae coli guided navigation and registration for virtual colonoscopy. IEEE Visualization, 279–285 (2005)
6. Huang, A., Roy, D., Summers, R.M., Franaszek, M., Petrick, N., Choi, J., Pickhardt, P.: Tenniae coli-based circumferential localization system for ct colonography: Feasibility study. Radiology 243, 551–560 (2007)
7. Lamy, J., Summers, R.M.: Teniæ coli detection from colon surface: Extraction of anatomical markers for virtual colonoscopy. In: Bebis, G., et al. (eds.) ISVC 2007, Part I. LNCS, vol. 4841, pp. 199–207. Springer, Heidelberg (2007)
8. Liang, Z., Wang, S.: An em approach to map solution of segmenting tissue mixtures: A numerical analysis, vol. 28, pp. 297–310 (2009)
9. Micha, S., Amir, S.: On shortest paths in polyhedral spaces. In: Proceedings of ACM Symposium on Theory of Computing, STOC 1984, pp. 144–153 (1984)
10. Mitchell, J.S.B., Mount, D.M., Papadimitriou, C.H.: The discrete geodesic problem. SIAM J. Comput. 16(4), 647–668 (1987)
11. O'Rourke, J.: Computational Geometry in C, 2nd edn. Cambridge University Press, New York (1998)
12. Summers, R.M., Yao, J., Pickhardt, P.J., Franaszek, M., Bitter, I., Brickman, D., Krishna, V., Choi, J.R.: Computed tomographic virtual colonoscopy computer-aided polyp detection in a screening population. Gastroenterology 129, 1832–1844 (2005)
13. Hampshire, T., Roth, H.R., Helbren, E., Plumb, A., Boone, D., Slabaugh, G., Halligan, S., Hawkes, D.J.: Endoluminal surface registration for ct colonography using haustral fold matching. Medical Image Analysis 17(8), 946–958 (2013)
14. Wan, M., Liang, Z., Ke, Q., Hong, L., Bitter, I., Kaufman, A.: Automatic centerline extraction for virtual colonoscopy. IEEE Transactions on Medical Imaging 21, 1450–1460 (2002)
15. 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)
16. Xin, S.Q., Wang, G.J.: Improving chen and han's algorithm on the discrete geodesic problem. ACM Trans. Graph. 28, 1–104 (2009)
17. Ge, Y., Stelts, D., Wang, J., Vining, D.: Computing the centerline of a colon: A robust and efficient method based on 3-d skeleton. Journal of Computer Assisted Tomography 23, 786–794 (1999)
18. Zeng, W., Marino, J., Chaitanya Gurijala, K., Gu, X., Kaufman, A.: Supine and prone colon registration using quasi-conformal mapping. IEEE Transactions on Visualization and Computer Graphics 16(6), 1348–1357 (2010)
19. Zhu, H., Barish, M., Pickhardt, P., Liang, Z.: Haustral fold segmentation with curvature-guided level set evolution. IEEE Transactions on Biomedical Engineering 60, 321–331 (2013)
20. Zhu, H., Li, L., Fan, Y., Lin, Q., Lu, H., Gu, X., Liang, Z.: Automatic teniae coli detection for computed tomography colonography. In: SPIE, Medical Imaging (2011)

Improved Colon Navigation for Efficient Polyp Detection in Virtual Colonoscopy

Marwa Ismail[1], Aly Farag[1], M. Sabry Hassouna[2], Gerald Dryden[3], and Robert Falk[4]

[1] CVIP Lab, University of Louisville, Brook St, Louisville, Kentucky, 40292
Marwa.ismail@louisville.edu
[2] Vital Images, Opus Pkwy , Minneapolis, USA, 55343
[3] Division of Gastroenterology/ Hepatology, E .Chestnut St, University of Louisville, USA, 40292
[4] Medical Imaging, Jewish Hospital, Abraham Flexner Way, Louisville, Kentucky, 40202

Abstract. Colon cancer is a leading cause of death in the world and its early diagnosis highly increases the chances of survival. Virtual colonoscopy is a widely spreading technology that is used for polyp detection, the primary cause of colon cancer. This paper revisits an existing virtual colonoscopy technique, called Fly-over. It splits the colon into two halves along its centerline and assigns a camera to each half for navigation. While cutting the colon along its centerline increases the possibility of having missed polyps, the technique is revisited here and the cutting framework is changed, which improved the rate of detection. Clinical validation was assessed by testing the navigation technique on several cases of real and synthetic challenging polyps versus other techniques. Fly-over technique provides efficient polyp detection of up to 100% with the least distortion rate.

Keywords: Colonoscopy, fly-through, fly-over, polyps, normal, distortion, centerline, rotation.

1 Introduction

Colorectal Cancer, cancer of the colon, rectum, anus, and appendix, is the third most commonly diagnosed cancer in the world. Most colorectal cancers begin as a polyp which is a superficial growth that arises from the colon wall, and as it grows, it can develop into a cancer that spreads [1]. Colonoscopy is the current gold standard for screening polyps, and although it detects more than 90% of the polyps, it is invasive and carries a risk of complications from sedation. These disadvantages introduced virtual colonoscopy (VC) which has evolved rapidly in the past decade and provides more comfort to patients, with significantly reduced side effects. It is not intended to replace standard colonoscopy, but rather will complement it by providing additional supportive information [2].

2 Related Work

The common visualization technique in VC is Fly- through (FT) navigation, where a virtual camera with a specified field of view moves along a special planned path

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 117–125, 2013.

inside the colon to render its internal views. The direction of navigation is either from the colon rectum side (antegrade) or from its cecum side (retrograde). The colon's centerline is agreed to be the optimal path that maximizes surface visibility during navigation [3]. The problem with FT is that up to 20% of the colon mucosa is not examined by a unidirectional navigation. The hidden zones are located mainly behind semicircular folds that become prominent when there is a colonic distension. A total of 4 fly-throughs should be performed (antegrade and retrograde in supine and prone positions) to examine the entire colon which is fatiguing and time- consuming [4].

Colon flattening methods were proposed to overcome FT problems. The main idea behind flattening is to initially transform the colon into a cylinder, and then map this cylinder onto a single image. Flattening methods introduce geometric distortion in general [5]. Panoramic view techniques aim at maximizing visible surface areas that are local to the camera's position along the centerline, yet polyps might be missed in some conditions [6].

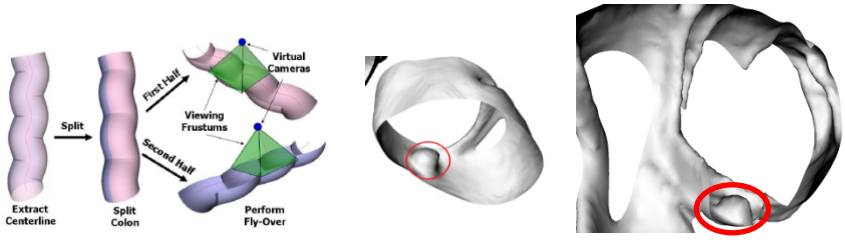
In [7], a visualization technique called Fly-over (FO) was proposed. The idea was to split the colon along its centerline into two halves, then assign a virtual camera to each half to perform FO navigation. While this method significantly improved the surface visibility coverage through only one navigation direction with the geometric distortion being significantly reduced, it still had some issues that are investigated in this paper. Cutting the colon along its centerline will generate a problem for polyps that the centerline passes through. This means that such polyps will be divided between the two halves, and thus lost in the detection process which is the ultimate goal of VC. Also, the method was not validated on a sufficient number of cases, such as detecting polyps in challenging areas (e.g. behind haustral folds), or at different anatomical positions. Also the degree of distortion of detected polyps was not assessed. This would've helped assessing the degree of loss of geometric properties of polyps with FO compared to the other techniques. These issues are covered in this paper.

Also to validate FO, the navigation path of FT was modified in an effort to maximize its visibility coverage of the colon walls and to help better validate FO. After this modification, FO was still preferred by physicians, after clinical validation on 30 sets, for some aspects including the reading time, and the distortion factor.

3 Methods

Initially, the colon lumen needs to get segmented, which is done by [8]. The centerline of the colon was extracted using [9], and then a 3D model of the colon was generated using the marching cubes algorithm [10]. In order to split the colon for FO navigation, the algorithm in [7] was adopted. Fig. 1(a) illustrates the FO technique. One of the major issues with this cutting framework is that polyps that pass through the centerline will be divided between the two halves, thus lost in the detection process (especially sessile

polyps that grow directly onto the colon wall, Fig. 2(b)). One way to handle this is to translate the centerline vertically upwards and downwards, then cut the colon with two planes along these two new versions of the centerline, instead of only one plane. Each division will thus have slightly more than half of the colon. The overlap between the 2 divisions that will occur as a result of this cutting guarantees that no polyps are missed, and if one polyp is cut in one division, it will be complete in the other. Cutting the colon with two planes will require calculating the normal vectors of these 2 planes for each group of the translated centerline points (upwards and downwards). The time required for this cutting framework is more than that of the original cutting by just few seconds. Equations that govern the new cutting framework are:



(a) (b) (c)

Fig. 1. a) Illustration of Fly-over technique, courtesy of [7]. b) A real polyp found in descending colon. c) A 10-mm synthesized one between haustral folds.

1-Generation of Rings. Assume the centerline C is composed of N points, and each point P_i is translated to $P_i \pm \Delta$. Each translated point will have its own tangential vector $\vec{t}_i$, and normal vector $\vec{n}_i$, which are given by:

$$\vec{t_\iota} = \frac{(P_{i+1} + \Delta) - (P_{i-1} + \Delta)}{2}, \tag{1}$$

$$\text{and} \quad \vec{n_\iota} = \vec{t_\iota} \times \vec{t_\iota} \times \vec{up}, \tag{2}$$

where $\vec{up}$ is an arbitrary vector that is chosen to be any of the Cartesian basis vectors. Equations (1) and (2) address translating the centerline upwards. The centerline is also divided into N_s segments using surface skinning. Points of each segment are assigned a specific label based on some label function:

$$\forall\, (P_i + \Delta) \in l_j, label(P_i + \Delta) = j\,. \tag{3}$$

Each mesh triangle is labeled by the label of the nearest centerline segment l_j and the set of surface triangles of the same label form one possible ring. If C_k is the geometric center of each surface triangle r_k, where $1 \le k \le N_r$, and N_r is the total number of surface triangles, then the skinning process is governed by:

$$\widehat{P + \Delta_\iota} = argmin\, \|C_k - (P_i + \Delta)\|^2, \tag{4}$$

$$label(r_k) = label\big(\widehat{P_i + \Delta}\big). \tag{5}$$

2- Splitting Rings. Each ring R_j has multiple centerline points. In order to split R_j into 2 halves, it is clipped using a plane π_j that is governed by:

$$\big(q - q_j\big).\overrightarrow{u_j} = 0, \tag{6}$$

where q_j is the starting point of the ring , q is a point that belongs to the plane, and $\overrightarrow{u_j}$ is its normal which is calculated by:

$$\overrightarrow{u_j} = \big(q_{j+1} - q_j\big) \times \overrightarrow{n_j}, \tag{7}$$

where q_{j+1} is the end point of the ring. The rings that form each part are concatenated altogether by a union operation to construct the whole colon. No distortion results because these rings complement each other.

Applying the same for centerline points that are translated downwards, we get the required two planes for cutting the colon into two parts, each contains more than half of it. Fig. 2 shows a sample result of the difference between the proposed cutting and the cutting in [7]. This minor change significantly improved the detection results as will be discussed in the next section.

4 Validation and Results

Navigation techniques were applied on 30 computed Tomography (CT) colonography sets. One has been provided by the 3DR Inc., Louisville, KY, and the rest were received from the Virtual Colonoscopy Center, Walter Reed Army Medical Center, Washington, DC. The patients underwent standard 24-hour colonic preparation by oral administration of 90 ml of sodium phosphate and 10 mg of bisacodyl; then consumed 500 ml of barium for solid-stool tagging and 120 ml of Gastrografin to opacify luminal fluid. The CT protocol included 1.25 to 2.5 mm collimation, and 100 mAs and 120 kVp scanner settings. Each dataset contains 400～500 slices, and the spatial resolution for is 1.0×1.0×1.0 mm^3.

In order to show improvements by the new cutting, several polyps were synthesized with medium to large sizes (10–30 mm), and implanted at different anatomical positions of the colon that were recommended by radiologists (e.g. rectum and sigmoid). Some polyps were implanted so that the centerline will pass through. Others were implanted at hard areas suggested by experienced radiologists and gastroenterologists (e.g. behind haustral folds, or in colon's curved parts, e.g. splenic flexure) to help quantitatively assess detection of polyps by FO versus other techniques.

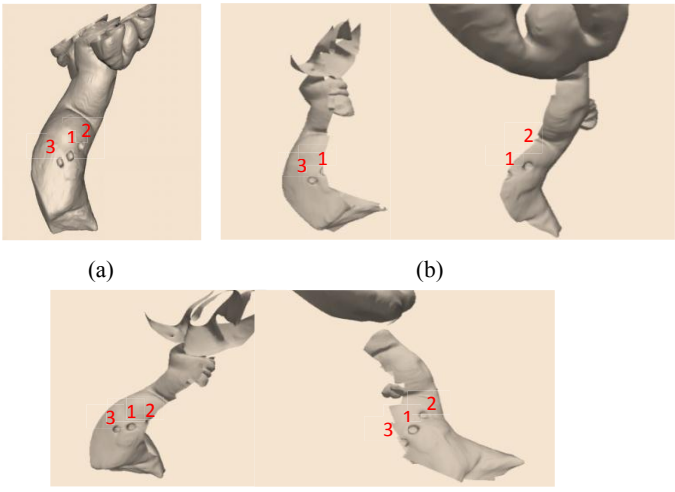


Fig. 2. a) Part of the colon (sigmoid), where 3 polyps were implanted, one that the centerline passes through (labeled by (1)), one is shifted to the right by 20 points (labeled by (2)), and one is shifted to the left by 15 points (labeled by (3)). b) Original Fly-over cutting, where polyp (1) is split as it passes through the centerline. c) The new cutting, where polyp (2) was missing in the first half, but found complete in the other and polyp (3) was missing in the second half but complete in the first.

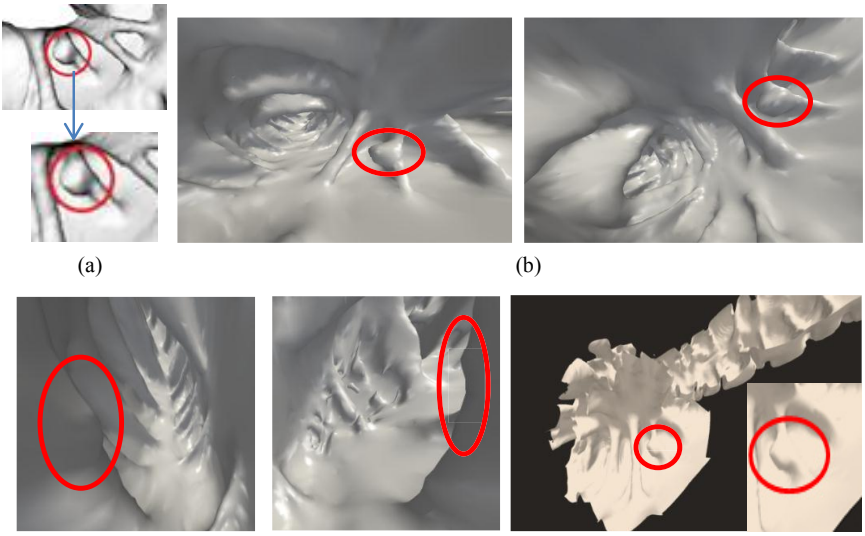


Fig. 3. a) A 10-mm synthesized polyp placed between folds. b) 2 sequences of the spiral motion of camera that could detect the polyp, but navigation time was long, and little distortion was present. c) Result of FT navigation in antegrade direction, where the polyp is severely distorted. d) The polyp was completely missed in retrograde navigation. e) FO navigation result of the same polyp easily detected as the direction of projection of the camera is perpendicular to the viewing volume. The geometry of the detected polyp is kept.

Colorectal polyps are usually found in one of two shapes: pedunculated; where polyps look like mushrooms, or sessile; where they grow directly onto the inner wall of the colon. Polyps could be synthesized as a group of intersected hemispheres, semiellipsoids, or a cylinder in union with a hemispherical cap. The inequality that governs the synthesis process is:

$$\frac{(x - x_0)^2}{a^2}\sin(\alpha) + \frac{(y - y_0)^2}{b^2}\cos(\alpha) + \frac{(z - z_0)^2}{c^2} \leq r^2, \tag{8}$$

where (x_0, y_0, z_0) is the center of the polyp, r is the desired radius, and (a, b, c) controls the shape whether it is a sphere, ellipsoid, or a cylinder. A transformation of coordinates is needed in order for the synthesized polyp to have the same local angle of the colon wall, which is provided by α. Intensity of polyps is set to be the same as that of the colon, and an averaging filter was applied to give the polyps smooth appearance. After inserting polyps based on equation (8) on the colon walls, their geometric features based on shape index (SI), curvedness (CV), and sphericity (SP) [11] were tested to validate the synthesis process. SI represents the topological shape of the object in the vicinity of a voxel, whereas CV provides information about how much of the object the neighborhood includes. SP ratio determines how much the shape is rounded. Equations that govern calculating geometric features can be found in [11]. According to [11], shapes of polyps belong to the cap class, where SI≅1. They also have small to medium CV values that range between 0.1 and 0.2. A perfectly rounded polyp has an SP value of 0, but it was relaxed here to be <1.2 for the elongated polyps. Synthesized polyps were designed to meet all of these criteria. Figs. 1(b) and 1(c) show a real polyp in the used sets and a synthesized one visualized in 3D. The sets used included 20 polyps that were read by 2 physicians independently. 25 other polyps were implanted, having a total of 45 polyps to be used in assessing the navigation techniques.

Two radiologists who are experienced on several colon navigation techniques reviewed the data sets independently in a blinded manner using FT, FO (original cutting), and modified FO. Any suspected polyps by the readers were to be confirmed by ground truth.

From experiments conducted on the 30 sets used, Fly-through proved to miss polyps at hard areas usually in one direction (either retrograde or antegrade). Moreover, some detected polyps are found to be severely distorted, Fig. 3(c). It is also time consuming to navigate the entire colon twice, i.e. (antegrade+retrograde) to search for polyps (colon is about 1.5 m long), with a high chance of getting distorted ones.

For efforts of getting better visibility coverage than this got with FT to provide a better comparison with FO (since FO is a clear winner over FT [7]), a modification was applied to its path. The navigation path is still going to be the centerline, but for each centerline point, the closest surface point to it on the colon wall is calculated, and will be the focal point of the corresponding centerline point. Before changing the camera position to the next centerline point, the focal point is radially rotated by 360^0 to cover the entire wall. For reasonable navigation time with better viewing results, the increment in the angle of rotation is 45^0, i.e. it spans one rotation in 8 steps before moving to the next centerline point. This provides a vertical rotation of the scene at each point and will enable to show details between haustral folds and reveal many spots that were blind to FT, which misses significant details, especially between folds.

This technique that uses spiral motion of the camera (we will call it spiral FT to be easily referred to) was also assessed by the two radiologists. It highly increased the rate of detection got by FT (Figs. 3(b) and 3(c)), yet increased the reading time considerably.

In FO where the old cutting was applied, most of the polyps the centerlines pass through were split and thus missed by the readers, especially the flat polyps where they were sometimes considered as folds, Fig. 4(d). The new cutting applied to FO technique, and although it is a minor tweak to the existing technology, helped detect almost all the existing polyps. Split polyps in one part could be found in the other. Figs. 3 and 4 show comparison results of the navigation techniques. Table 1 shows quantitative results of comparing them in terms of the colon surface visibility coverage (determined using the Z-Buffer visibility test), the percentage of detected polyps, the degree of distortion; measured as the ratio between shortest and longest diameter of a polyp appearing on the screen by each navigation technique [6], and navigation time required, which represents the average reading time by the two involved physicians for all the data sets. Degree of distortion for perfectly spherical polyps should be one, and as most of the polyps found in sets (real or synthesized) were close to spherical or hemispherical in shape, the degree of distortion should be close to this value. This indicates that the more the polyp is distorted, the greater the distortion value will be (>1). Sensitivity and specificity of polyp detection by each technique were also calculated. It is clear that the new cutting significantly increased the overall percentage of polyp detection with minimal distortion. Although the spiral FT proved to be

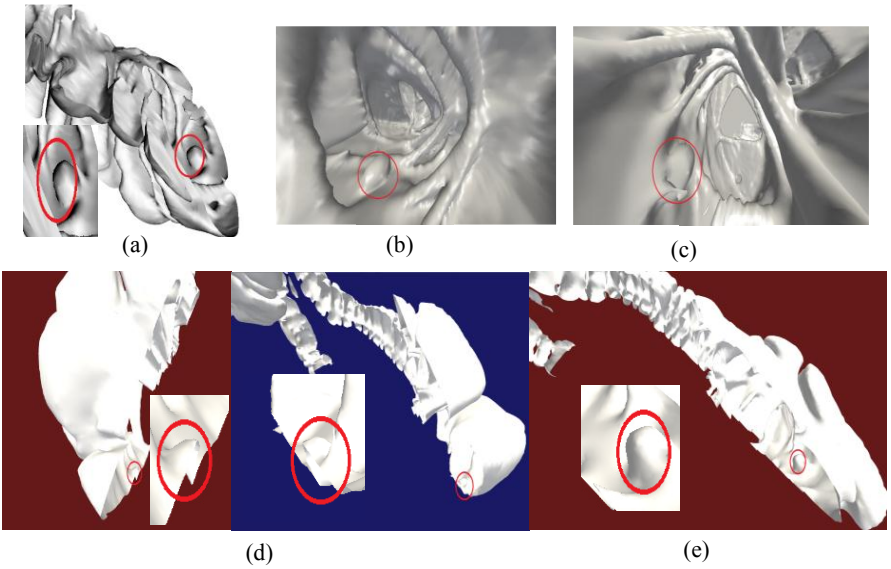


Fig. 4. a) A 12- mm synthetic polyp designed to pass through the centerline of the colon. b) The polyp is distorted by 50% in FT antegrade navigation. c) FT retrograde navigation with about 70% distortion. d) Original FO cutting, where the polyp is split into the 2 halves (red and blue). Part of the split polyp in the blue half is hidden by a fold. e) The polyp is complete and kept its geometry in the proposed cutting.

highly efficient in detecting polyps due to the rotation of each focal point that allows spanning the entire colonic wall, FO was still preferred by physicians as it easily readable, less distorting, and its reading time for a procedure is much shorter, which is a key in VC.

Table 1. Comparison results of virtual colonoscopy navigation techniques

Technique		FT antegrade	FT retrograde	FT (antegrade +retrograde)	Original FO	Improved FO	Spiral FT
% of detected polyps	Through Centerline	70%	70%	75%	88%	95%	93%
	Hard areas	72%	70%	75%	92%	92%	90%
	Total	70%	70%	75%	90%	94%	91%
Visibility Coverage		79%	78%	93%	99%	99%	95%
% of Distortion		50%	70%	65%	3%	3%	20%
Navigation Time (min) for 1 set		10	10	16	6	6-7	20
Sensitivity		70%	70%	82%	90%	99%	95%
Specificity		70%	70%	80%	92%	99.5%	95%

5 Conclusions

Virtual Colonoscopy is a non-invasive, and a comfortable procedure for screening polyps, the main cause of colon cancer. In this paper, virtual Fly-over technique was modified to enhance its chances for polyp detection. Also while validating the work, a modification was done to the navigation path of FT to increase its visibility coverage in order to be a good competitive to FO. Real and synthetic polyps in challenging positions were introduced to validate this improvement versus the original technique and versus Fly-through technique. FO proved to be a winner in many aspects, including ease of readability, navigation time, distortion index, and rate of polyp detection. Future work will focus on incorporating FO with flattening techniques. Flattening techniques introduce notable distortion in general, but are preferred by a wide number of physicians due to its ease of readability. If FO is applied after the colon is flattened, it will introduce a significant improvement per polyp detection rate, where the suspected findings of polyps from the flattened colon can be checked using FO.

References

1. Abbruzzese, J., Pollock, R.: Gastrointestinal Cancer. Springer (2004)
2. Buthiau, D., Khayat, D.: Virtual Endoscopy. Springer (2003)
3. Bouix, S., Siddiqi, K., Tannenbaum, A.: Flux Driven Fly Throughs. In: Proc. of IEEE CVPR 2003, pp. 449–454 (2003)
4. Pickhardt, P.J., Taylor, A.J., Gopal, D.V.: Surface Visualization at 3D Endoluminal CT Colonography: Degree of Coverage and Implications for Polyp Detection. Gastroenterology 130, 1582–1587 (2006)
5. Hong, W., Gu, X., Qiu, F., Jin, M., Kaufman, A.: Conformal Virtual Colon Flattening. In: Proc. of ACM Symposium on SPM 2006, pp. 85–93 (2006)
6. Lee, S.S., Park, S.H., Kim, J.K., Kim, N., Lee, J., Park, B.J., Kim, Y.J., Kim, A.Y., Ha, H.K.: Panoramic Endoluminal Display with Minimal Image Distortion Using Circumferential Radial Ray-Casting for Primary Three-Dimensional Interpretation of CT Colonography. Eur. Radiol. 19, 1951–1959 (2009)
7. Hassouna, M.S., Farag, A.A., Falk, R.: Virtual Fly-Over: A New Visualization Technique for Virtual Colonoscopy. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 381–388. Springer, Heidelberg (2006)
8. Ismail, M., Elhabian, S., Farag, A., Dryden, G., Seow, A.: Fully Automated 3D Colon Segmentation for Early Detection of Colorectal Cancer Based on Convex Formulation of the Active Contour Model. In: Proc. of IEEE CVPRW 2012, pp. 58–63 (2012)
9. Hassouna, M.S., Farag, A.A.: Accurate Tracking of Monotonically Advanced Fronts. In: Proc. of IEEE CVPR 2006, pp. 355–362 (2006)
10. Lorensen, W.E., Cline, H.E.: Marching Cubes: a High Resolution 3D Surface Construction Algorithm. SIGGRAPH Comput. Graph. 21, 163–169 (1987)
11. Ismail, M., Elhabian, S., Farag, A., Dryden, G., Seow, A.: 3D Automated Colon Segmentation for Efficient Polyp Detection. In: Proc. of Cairo International Conference on Biomedical Engineering (CIBEC), pp. 48–51 (2012)

Personalised Estimation of the Arterial Input Function for Improved Pharmacokinetic Modelling of Colorectal Cancer Using dceMRI

Benjamin Irving[1], Lydia Tanner[1], Monica Enescu[1], Manav Bhushan[1], Esme J. Hill[3], Jamie Franklin[2], Ewan M. Anderson[2], Ricky A. Sharma[3], Julia A. Schnabel[1], and Michael Brady[3]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Old Road Campus Research Building, Oxford, UK, OX3 7DQ
benjamin.irving@eng.ox.ac.uk
[2] Department of Radiology, Churchill Hospital, Old Road, Oxford, UK, OX3 7LE
[3] Department of Oncology, University of Oxford, Old Road Campus Research Building, Oxford, UK, OX3 7DQ

Abstract. dceMRI is becoming a key modality for tumour characterisation and monitoring of response to therapy, because of the ability to identify the underlying tumour physiology. Pharmacokinetic (PK) models relate the contrast enhancement seen in dceMRI to physiological parameters but require accurate measurement of the AIF, the time-dependant contrast concentration in blood plasma. In this study, a novel method is introduced that overcomes the challenges of direct AIF measurement, by automatically estimating the AIF from the tumour tissue. This approach was evaluated on synthetic data (10% noise) and achieved a relative error in K^{trans} and k_{ep} of $11.8 \pm 3.5\%$ and $25.7 \pm 4.7\%$, respectively, compared to $41 \pm 15\%$ and $60 \pm 32\%$ using a population model. The method improved the fit of the PK model to clinical colorectal cancer cases, was stable for independent regions in the tumour, and showed improved localisation of the PK parameters. This demonstrates that personalised AIF estimation can lead to more accurate PK modelling.

Keywords: pharmacokinetic modelling, arterial input function, dceMRI.

1 Introduction

Dynamic contrast-enhanced magnetic resonance imaging (dceMRI) is becoming increasingly common for monitoring cancer response to therapy because of its ability to identify the underlying tissue physiology, such as microvasculature and capillary leakage, from tissue contrast agent (CA) uptake. During dceMRI acquisition, a bolus injection of CA is injected into a peripheral vein, which travels through the vascular system and leaks from the capillary network into the tissue extravascular-extracellular space (EES). Re-uptake and renal excretion then lowers the CA. This observed signal enhancement (S_e) can be modelled as

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 126–135, 2013.

a convolution of an arterial input function (AIF) with a pharmacokinetic (PK) model to extract the physiological parameters of the tumour.

Accurate measurement of the AIF (the CA concentration in blood) is required for calculation of the AIF and PK convolution, however direct measurement from an artery is invasive and adds complexity to the dceMRI acquisition. The AIF can be measured directly from the dceMRI image by identification of an artery [1,2], but the temporal resolution of dceMRI is generally too low to capture the initial AIF shape. Including an additional perfusion CT scan has been proposed but is often not feasible and adds radiation dose [3]. Population models of the AIF may be used as a substitute for direct AIF measurements [4,2]. However, there is considerable inter- and intra-patient variability due to a number of physiological factors including heart rate, renal function and injection timing. We propose a novel approach to estimate the AIF from the tissue concentration curves, which does not require use of any additional measurements, modalities or reference regions. The method improves upon the initial population AIF estimate by jointly determining the optimal patient specific AIF and PK parameters from the tumour tissue region of interest (ROI). Unlike attempting to directly measure the AIF from an artery in dceMRI, a high temporal resolution is not required when estimating AIF from tissue concentration.

Some previous studies have developed methods to estimate patient specific AIFs. Liberman *et al.* [5] perform a search for optimal AIF parameters but require strict selection of voxels (brain grey and white matter); concentration curves with a visible proportion of plasma, which may not be present especially with a low temporal resolution; and find AIF and PK parameters independently – not accounting for interdependence. A similar method to ours is presented by Fluckiger *et al.* [6], by iteratively fitting the Tofts model [7] and an AIF model to eight representative curves derived from the tissue ROI. However, the method requires an AIF measurement to normalise the model; computationally intensive calculation of the convolution for each curve – therefore limited to eight curves; and an AIF model that requires 11 parameters to be fitted. The method we propose overcomes these limitations by including: knowledge of the population mean and variation to initialise and constrain the model, and an AIF model that requires fewer parameters. This allows an analytic solution to be found for the tissue concentration curve, which in turn speeds up the optimisation and allows the AIF to be optimised directly on at least 500 voxel tissue CA curves. No additional AIF measurements are required. Section 2 outlines our patient specific AIF estimation method to improve PK modelling, synthetic and clinical datasets are introduced in Section 3 and used for evaluation in Section 4.

2 Method

This section outlines our method to jointly determine the patient-specific AIF and PK parameters for a region of interest using dceMRI. As shown in Figure 1, the population AIF and tumour $S_e(t)$ curves are input into the model (Section 2.2). In this study, the Orton AIF [4] with Tofts PK model [7] (Section 2.1) are

used. However, other choices could be incorporated into this general framework. This model is fitted to each input $S_e(t)$ to derive the PK parameters. These parameters are then held fixed while a nonlinear fit determines the optimal AIF for the entire region. The updated AIF is used to generate an improved PK model fit. This is repeated using an alternating minimisation method until a final AIF and set of PK parameters for the region is found (Section 2.3).

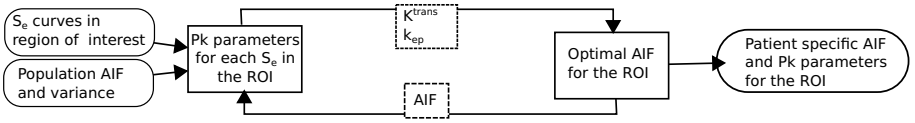


Fig. 1. The alternating minimisation framework for determination of patient specific AIF and improved pharmacokinetic parameter calculation

2.1 Tofts-Kety Model with Orton AIF

A common PK model is the Tofts-Kety two-compartment model where contrast is transferred between the blood plasma compartment and EES compartment [7]. Tissue concentration is described by $\frac{dC_e}{dt} = K^{trans}C_p - k_{ep}C_e$ with solution

$$C_e(t) = K^{trans}C_p(t) \otimes e^{-k_{ep}t} \tag{1}$$

where $\otimes$ is the convolution operator and change in tissue CA concentration (C_e) is determined by the CA concentration in blood plasma (C_p) and current C_e [7,4]. As can be seen, C_e is dependant on accurate determination of the C_p (or AIF). This model can be fitted to concentration curves in dceMRI to derive physiological parameters (K^{trans}, k_{ep} and $v_e = K^{trans}/k_{ep}$). In this study we have made the assumption that the plasma fraction in the tumour is negligible ($v_p \approx 0$) [4]. The plasma fraction makes a contribution at the early vascular uptake stage which requires a high temporal resolution to detect. Bradley *et al* [8] also suggest that colorectal tumours have a low plasma fraction.

Orton *et al.* [4] provide a population model of the AIF that allows the Tofts model to be explicitly solved for C_e. The analytic solution to the Tofts-Kety model (Equation 1) with Orton AIF is:

$$\begin{aligned}
C_e(t) =& \frac{A_1A_2K^{trans}}{k_{ep}-m_2}\Big(f(t,m_2) + f(t,k_{ep})\Big(\frac{k_{ep}-m_2}{A_2} - 1\Big)\Big) \text{ for } 0 \le t \le t_B \\
C_e(t) =& \frac{A_1A_2K^{trans}}{k_{ep}-m_2}\Big(f(t_B,m_2)e^{-m_2(t-t_B)} \\
&+ \Big(\frac{k_{ep}-m_2}{A_2} - 1\Big)f(t,k_{ep})e^{-k_{ep}(t-t_B)}\Big) \text{ for } t > t_B \qquad (2) \\
\text{with } f(t,\alpha) =& \frac{1}{\alpha}(1-e^{-\alpha t}) - \frac{1}{\alpha^2+m_1^2}(\alpha\cos(m_1t) + m_1\sin(m_1t) - \alpha e^{-\alpha t})
\end{aligned}$$

where A_1, A_1, m_1 and m_2 are the parameters of the AIF, K^{trans} and k_{ep} are the PK parameters of interest, and $t_B = 2\pi/m_1$. AIF offset (τ) is included in the AIF model to allow the bolus start time to be automatically determined such that $t = \hat{t} - \tau$ and $C_e(t) = 0$ for $t < 0$, where $\hat{t}$ is the measured time.

The relationship between tissue concentration $C_e(t)$ and observed signal enhancement $S_e(t)$ in dceMRI is nonlinear. $C_e(t)$ was converted to $S_e(t)$, using $S_e = \exp(-r_2 C(t) T_E \frac{1-\exp(-P-Q)-\cos\alpha(\exp(-P)-\exp(-2P-Q))}{1-\exp(-P)-\cos\alpha(\exp(-P-Q)-\exp(-2P-Q))}$, where $P = T_R/T_{10}$ and $Q = r_1 C(t) T_R$, TR is the repetition time, TE is the echo time and r_1 and r_2 are known constants and α is the flip angle of the dceMRI image, from the Spoiled Gradient Recalled (SPGR) sequence model [9].

A key limitation of fitting Equation 2 is that globally scaling the AIF parameter A_1 over the entire region has the same result as scaling K^{trans} for each voxel. Therefore, A_1 is fixed as the population average to form an additional constraint on the AIF model. In our method, we use this analytic solution to derive the patient specific AIF for the ROI and voxelwise PK parameters.

2.2 Population AIF

In our AIF estimation method, the population AIF with variance is used to initialise and constrain the optimisation. Parker *et al.* [2] measured patient AIFs directly from high temporal-resolution dceMRIs of arteries of 67 scans of patients between 18 - 80 years. They used these measurements to develop a parametric function of AIF variation in a population. This model cannot be used to solve C_e analytically, and was used in our study to generate 1000 arterial input functions by sampling from the population distribution. These curves provide an estimate of the population mean and variance, and Orton *et al.*'s [4] model was fitted to each curve to generate the population mean and standard deviation for the model parameters: $A_1 = 2.65 \pm 0.18 mmol/l$, $A_2 = 1.51 \pm 0.68 mmol/l$, $m_1 = 22.40 \pm 0.73 min^{-1}$ and $m_2 = 0.23 \pm 0.46 min^{-1}$. These mean values are used as the initial AIF in order to derive a patient specific AIF.

2.3 Alternating Minimisation Method and Constraints

The previous sections outline the AIF and PK models that are used in our patient specific AIF estimation method. Optimal parameters for both the models (described in Section 2.1) are found from the dceMRI ROI in an alternating minimisation method. The alternating minimisation method is initialised using the population AIF, and the model is fitted to each dceMRI S_e curve to derive the PK parameters using the 'trust-region-reflective' method for non-linear least squares curve fitting. In this initial step, PK parameters are derived for each voxel. These PK parameters are used to fit the same model over the entire region to determine the optimal AIF parameters, while keeping the PK parameters fixed. This new AIF model is then used to improve the individual PK parameter estimation for each voxel. This is repeated until convergence (shown in Figure 1).

Bounds are set on the AIF parameters to be within 4 standard deviations of the population average – representing 99.99% of the population – and a bound of $v_e \leq 1$. The lumen may be included in the ROI due to movement from peristalsis. To exclude these voxels, voxels that show no enhancement at temporal postions after initial enhancement are excluded.

Therefore, our model is able to obtain patient specific AIF and PK parameters without additional blood sampling, scans or the assumption of a population AIF. This method was evaluated on synthetic and clinical datasets.

3 Materials

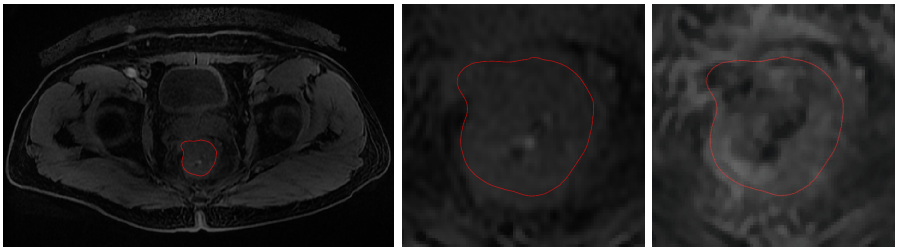
3.1 Synthetic Images

Synthetic data was used to evaluate the method by generating signal enhancement (S_e) curves from known PK parameters and AIFs. These known PK parameters and AIFs were compared to the parameters derived by our method on the synthetic data. 30 synthetic cases were generated, each consisting 500 signal enhancement curves. These were generated with 0%, 5% and 10% added noise. Each of the 500 voxels consisted of 29 temporal positions with a temporal resolution of 9.5s – consistent with the clinical data in this Section.

Each case consisting of 500 voxels was generated from the same sampled AIF and each voxel from a set of PK parameters. AIFs were derived for each case by sampling from a Gaussian distribution over the means and standard deviations of the Orton AIF parameters. This AIF was kept fixed for each of the 10 datasets while generating 500 K_{trans} and k_{ep} pair to calculate individual signal enhancement curves. K_{trans} was derived from a uniform distribution on the interval [0,2] and v_e was derived from [0.4,1], $k_{ep} = K_{trans}/v_e$, and the offset was sampled from [0, 0.5] minutes. These parameters and the AIF values were then used with the Tofts and Orton AIF model to generate the 500 signal enhancement curves. Noise was added to each temporal position of each voxel by sampling from a Gaussian distribution with standard deviation equal to 0%, 5% and 10% of the mean enhancement of the curve. Example synthetic S_e curves are shown in Figure 5c) and Figure 5f).

3.2 dceMRI Colorectal Cancer Cases

Six dceMRI images were acquired as part of a phase 0/1 drug trial with hypofractionated radiotherapy in patients with colorectal cancer. Pre- and post-treatment image sequences were acquired. A 3T GE scanner was used to acquire the dceMRI images with the LAVA protocol and spoiled gradient echo sequence. Images with dimensions 512x512x52x29 were acquired with voxel size of 0.78x0.78x2.00mm, and a temporal resolution of 9.5s. ProHance (Gadoteriol) contrast was injected at a rate of 3 ml/sec, 0.1 mmol/kg body weight. Colorectal tumours were delineated by a clinician on T2 weighted images and registered to the dceMRI image. Flip-angle images were not available and a uniform T_{10} map of 1 was assumed. Figure 2 shows a cross section of a tumour ROI before and after CA enhancement. K^{trans} maps of this ROI are shown later in Figure 7.



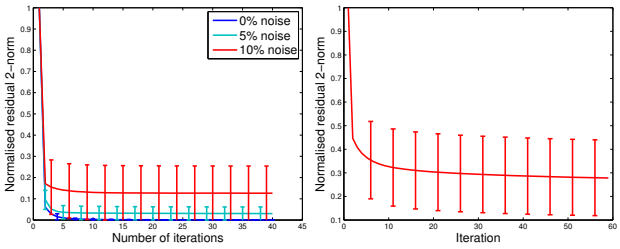
(a) (b) (c)

Fig. 2. dceMRI slice through centre of colorectal tumour for Patient 1, with b) zoomed in ROI before contrast (0s) and c) after contrast enhancement (276s)

4 Results

4.1 Results Using Simulated Data

Our patient specific AIF estimation method was evaluated on each of the 10 sets of data at each noise level of 0%, 5% and 10%. The alternating minimisation method was run for 100 iterations to determine the patient specific PK parameters and AIF. Figure 3 shows the normalised residual 2-norm error in the fit of the model to the signal enhancement curves for a) the first on synthetic data and b) the on clinical data with the standard deviation representing the variation between the cases. The first iteration is the fit using a population AIF and the remaining iterations show the error as a proportion of this initial fit. As expected, there is a considerable improvement in fit of the pharmacokinetic model to the signal enhancement curves by finding a patient specific AIF. Synthetic data at 10% noise and clinical data achieved fits with errors of $12.6 \pm 12.8\%$ and $27.8 \pm 16.0\%$, respectively, of the initial population model fit.



(a) Synthetic cases (b) Clinical cases

Fig. 3. Normalised residual 2-norm error for each iteration of our AIF estimation method a) synthic cases with 0%, 5% and 10% added Gaussian noise b) clinical cases. These results are normalised by the error using the population AIF (first iteration).

The relative error and correlation between the derived PK parameters and the known PK parameters (used to generated the synthetic data) are shown in Figure 4. Our patient specific AIF method achieves a very high correlation between the 'truth' and derived PK parameters –important for analysis of heterogeneity in tumours. The accuracy of the PK parameters is also considerably improved compared to the PK parameters derived using a population AIF. K^{trans} relative error with 10% noise was $11.8 \pm 3.5\%$ (from $41 \pm 15\%$) and k_{ep} error of 25.7 ± 4.7 (from $60 \pm 32\%$). This demonstrates the benefit of our method and the importance of determining a patient specific AIF for PK parameter calculations. Noise impacts the accuracy of derived PK parameters (particularly k_{ep}) but are still considerably better than the population AIF derived values even without noise. There is a small error in K^{trans} even without added noise due to the assumption of a population A_1 parameter.

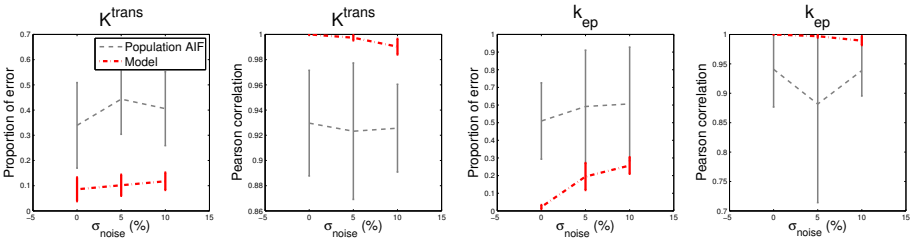


Fig. 4. Error and correlation of K^{trans} and k_{ep} using the population average and patient specific AIF on synthetic cases. The dashed grey lines show the calculated K^{trans} and k_{ep} using the population AIF while the red lines show the PK parameters calculated using the derived patient specific AIF. The x-axis is the standard deviation of the Gaussian noise in the synthetic data and the error bars show the standard deviation of the relative error in the 10 datasets.

Two examples of the AIF optimisation are shown in Figure 5a-c) and Figure 5d-f). The derived patient specific AIFs matched the known AIFs closely. In Figure 5a-c), the PK parameters used to generate the synthetic curve were K^{trans}=0.608 and k_{ep}=0.615. Parameters derived using the population AIF were less accurate (0.712 and 0.864) compared to those derived using the patient specific AIF (0.548 and 0.615). In this example, there is little offset or bolus peak difference between the population AIF and 'true' AIF. However, just the increased tail of the AIF has a considerable impact on the concentration curve fit and accuracy of PK parameters, particularly k_{ep}.

4.2 Results Using Colorectal dceMRI

Our patient specific AIF method was applied to six clinical cases. Figure 3b) shows that a considerably improved fit to the SE curves in the scan is achieved using a patient specific AIF ($< 30\%$ of the original error). Therefore, the patient

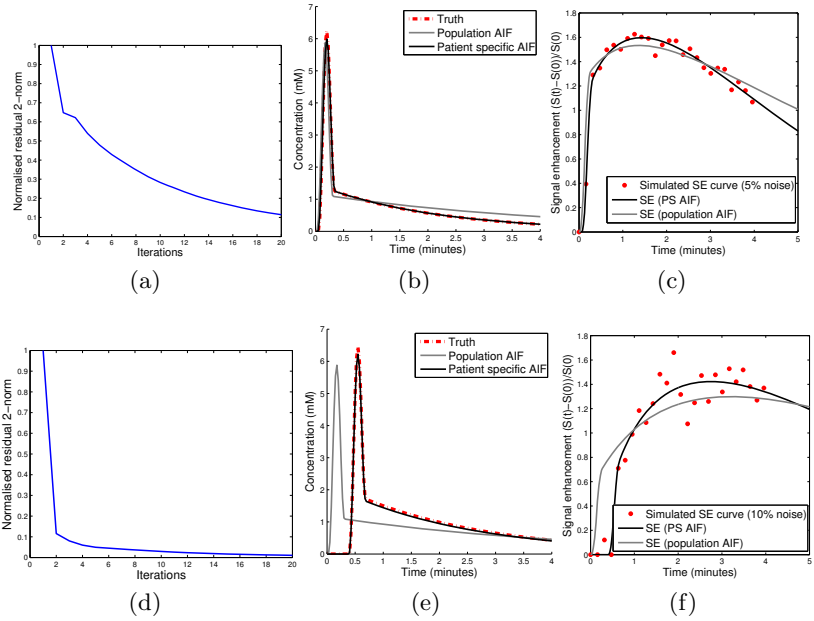


Fig. 5. Patient specific AIF compared to the known AIF for two synthetic cases a-c and d-f. a) and d) show the error between the known AIF and the estimated AIF over the first 20 iterations of the algorithm. b) and e) compare the population AIF and the derived patient specfic AIF to the known AIF used to generate the synthetic data (dashed curve). c) and e) show one example of the 500 signal enhancement curves in a case that are used to derive the parameters, where grey is the fitted curve using the population AIF and black is the fitted curve using our patient specific AIF.

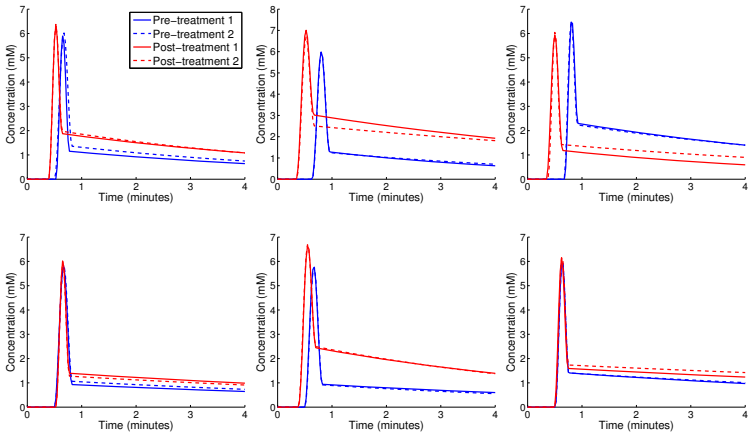


Fig. 6. Pre- and post-treatment AIFs found for 6 clinical colorectal cancer cases. Two AIFs were generated using independent regions in the tumour in each image.

specific AIF results in a PK model that is more representative of tumour signal enhancement. Figure 6 shows AIF curves generated for pre-treatment (gray) and post-treatment (black) for the centre tumour slice (solid line) and the following slice (dashed lines) for 6 cases. The similarity between the AIF derived from neighbouring slices shows that the model is stable when optimised on independent regions of the tumour. Interestingly, the AIFs show a trend of having a larger 'tail' for post therapy images. Therefore, as well as improving the accuracy of PK parameter estimation, AIF shape may have diagnostic value. Figure 7 shows example K^{trans} maps for the centre slice of the tumour ROI for the pre and post radiotherapy using both the population and derived patient specific AIF. Increased localisation and defined focal points of activity are shown with use of the correct offset and patient specific AIF.

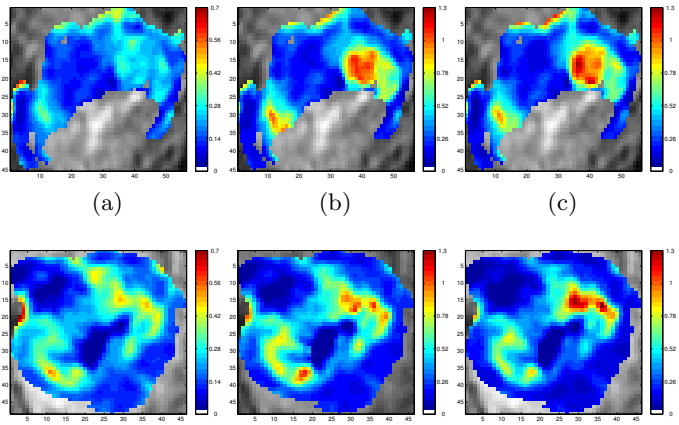


Fig. 7. K^{trans} maps for a cross-section of the tumour ROI (color) with image background for Patient 1. a)-c) show pre-radiotherapy and d-f) post radiotherapy. a) and d) Population AIFs with an incorrect offset (30s error), b) and e) Population AIFs with corrected offset and c) and f) patient specific AIFs derived using our model.

5 Discussion and Conclusions

PK parameter calculation in dceMRI requires accurate measurement of the AIF. This is often not possible and a population AIF is used instead, which results in inaccuracies. In this study, a novel method to obtain a patient specific AIF from the population model and tissue ROI is introduced. Unlike previous methods, our method does not require additional scans, direct measurements or identification of an artery. The optimisation is also applied directly to a large number of signal enhancement curves in the ROI, making the fit robust and adaptable. This method considerably improves the PK model fit compared to the population

AIF for both synthetic and clinical dceMRI cases, leading to more accurate PK parameters. In synthetic data, the K^{trans} relative error with 10% noise was $11.8 \pm 3.5\%$ (from $41 \pm 15\%$) and k_{ep} error of 25.7 ± 4.7 (from $60 \pm 32\%$). In clinical cases the AIF produces robust results for independent regions in the tumour. Future work will include: incorporation of motion correction into the method and use of this patient specific AIF model to better estimate patient response to therapy. There is also potential to compare the personalised AIF to direct AIF measurements from high temporal resolution scans of an artery, and also examine the effect of temporal resolution on AIF estimation.

Acknowledgements. This work was supported by the CRUK/EPSRC Oxford Cancer Imaging Centre. We thank Prof. Fergus Gleeson for the dceMRI dataset.

References

1. Rijpkema, M., Kaanders, J.H., Joosten, F.B., van der Kogel, A.J., Heerschap, A.: Method for quantitative mapping of dynamic MRI contrast agent uptake in human tumors. Magn. Reson. Imaging 14, 457–463 (2001)
2. Parker, G., Roberts, C., Macdonald, A., Buonaccorsi, G., Cheung, S., Buckley, D., Jackson, A., Watson, Y., Davies, K., Jayson, G.: Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI. Magn. Reson. Med. 56, 993–1000 (2006)
3. Enescu, M., Bhushan, M., Hill, E.J., Franklin, J., Anderson, E.M., Sharma, R.A., Schnabel, J.A.: pCT derived arterial input function for improved pharmacokinetic analysis of longitudinal dceMRI for colorectal cancer. In: SPIE, p. 86690 (2013)
4. Orton, M.R., D'Arcy, J.A., Walker-Samuel, S., Hawkes, D.J., Atkinson, D., Collins, D.J., Leach, M.O.: Computationally efficient vascular input function models for quantitative kinetic modelling using DCE-MRI. Phys. Med. Biol. 53, 1225–1239 (2008)
5. Liberman, G., Louzoun, Y., Colliot, O., Ben Bashat, D.: T_1 Aapping, AIF and Pharmacokinetic Parameter Extraction from Dynamic Contrast Enhancement MRI Data. In: Liu, T., Shen, D., Ibanez, L., Tao, X. (eds.) MBIA 2011. LNCS, vol. 7012, pp. 76–83. Springer, Heidelberg (2011)
6. Fluckiger, J.U., Schabel, M.C., DiBella, E.V.R.: Model-based blind estimation of kinetic parameters in dynamic contrast enhanced (DCE)-MRI. Magn. Reson. Med. 62, 1477–1486 (2009)
7. Tofts, P.S., Brix, G., Buckley, D.L., Evelhoch, J.L., Henderson, E., Knopp, M.V., Larsson, H.B.W., Lee, T.Y., Mayr, N.A., et al.: Parker: Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols. Magn. Reson. Imaging 10, 223–232 (1999)
8. Bradley, D.P., Tessier, J.L., Checkley, D., Kuribayashi, H., Waterton, J.C., Kendrew, J., Wedge, S.R.: Effects of AZD2171 and vandetanib (ZD6474, Zactima) on haemodynamic variables in an SW620 human colon tumour model: an investigation using dynamic contrast-enhanced MRI and the rapid clearance blood pool contrast agent, vol. 20, pp. 42–52 (2008)
9. Tofts, P.S., Kermode, A.G.: Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts. Magn. Reson. Med. 17, 357–367 (1991)

Free-Form Registration Involving Disappearing Structures: Application to Brachytherapy MRI

Floris F. Berendsen[1], Alexis N.T.J. Kotte[2], Astrid A.C. de Leeuw[2], Max A. Viergever[1], and Josien P.W. Pluim[1]

[1] Image Sciences Institute, University Medical Center Utrecht, P.O. Box 85500, 3508 GA, Utrecht, The Netherlands
floris@isi.uu.nl

[2] Department of Radiotherapy, University Medical Center Utrecht, P.O. Box 85500, 3508 GA, Utrecht, The Netherlands

Abstract. Registration of two images is difficult if large deformations are induced due to the absence of a structure in one image. We propose a penalty term that minimizes the volume of the missing structure in one image during free-form registration. The registration optimum found is based on image similarity, provided that the missing volume is minimal. We demonstrate our method on cervical MR images for brachytherapy. The intrapatient registration problem involves one image in which a therapy applicator is present and one in which it is not. Experiments show improvement of registration when including the penalty term. The improvements of surface distance and overlap of the bladder and rectum (which are close to the applicator volume) provide proof of principle of our method.

Keywords: registration, regularization, missing correspondence, surface mesh.

1 Introduction

Registration algorithms generally assume there exists a one-to-one mapping between images, e.g. for every region in one image a corresponding region can be found in the other. However, in several applications this assumption is violated, for instance after tissue resection or bone drillout or after withdrawal of a brachy applicator. The missing volume can affect the anatomy by a either (or both) of these effects [1]:

(i) The missing volume leaves a void (filled with air or fluid);
(ii) The surrounding soft tissue fills in the missing volume.

Most approaches described in the literature address (i), with brain resection as the primary application. Since (manual) masks of the missing volume can be absent [2], [3], [4] or only probabilistic [5], [1], frameworks that perform segmentation and registration simultaneously are proposed. Correctly retrieving the regions that have correspondence is crucial for reducing spurious deformation at

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 136–144, 2013.

the resection edges. Approaches that do account for (ii) incorporate a tumor model [6] or use finite element approaches [7],[8] to model brain deformation.

In this paper we focus on the type (ii) effect, proposing a method that is not based on a model, because modeling can be problematic in some applications. Our application is on registration of cervical MR images. For the treatment of cervical cancer a combination of external beam radiotherapy and (internal) brachytherapy is often used. Fig. 1 shows an example treatment applicator. Before either treatment the patient is imaged (using MR) for target delineation.

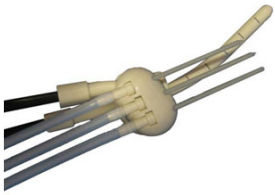


Fig. 1. The Utrecht Interstitial CT/MR Applicator (Nucletron, Veenendaal)

To investigate the amount of dose that the combined therapies deliver to specific tissue, the correspondence between the images of both therapies needs to be found. Registration is challenging because a treatment applicator is present in the brachytherapy image and not in the image of external therapy. This missing volume is not only large, its absence also has a big influence on the topology of the surrounding tissue.

To cope with this we propose a penalty term that minimizes the volume of the applicator void in a general non-rigid registration framework. By including this prior knowledge, we expect better registration, because the final shape the void collapses into will be driven by the image data, subjected to minimization of its volume.

2 Method

As registration framework we use the elastix registration toolbox [9]. The segmentation of the missing volume is assumed to be known in one image and its contour is represented by a surface mesh.

The cost function in the registration scheme consists of an image similarity measure $\mathcal{S}$ and surface mesh penalty function $\mathcal{P}$:

$$\mathcal{C} = -\mathcal{S} + \alpha\mathcal{P} \ , \tag{1}$$

where α balances the two terms. The image similarity is calculated for the images where the region of the disappearing structure (the applicator) is masked out in one image. One transformation model $T_\mu(\mathbf{x}) : \Omega_{\text{fixed}} \mapsto \Omega_{\text{moving}}$ applies to both the image and the surface mesh and is represented by uniform cubic B-splines

with coefficients μ. A uniform B-spline transformation model is used since it allows efficient implementation.

The image containing the volume that will be minimized (the applicator) is chosen as fixed image, the other as moving image. Only this definition of domains allows a (surjective) mapping of a certain volume to zero to be modeled by uniform B-splines. We specifically choose not to force all Jacobian determinants within the applicator volume to be zero, since this would require a great number of costly evaluations. Additionally, having a region that is fully compressed to zero would have a negative impact on the deformations in neighboring regions, since the transformation model enforces smoothness. Instead, a penalty function that only acts on the boundary of the applicator void is designed such that a minimal net volume is imposed. As a consequence, the transformations inside the applicator region can exhibit a lot of folding, but this does not need to be resolved, because the region is masked out for registration.

Our missing structure penalty term (MSP) was designed to allow computation of an analytical derivative (for gradient-based optimizers) and to prevent self intersection of the boundary.

2.1 Volume Penalty

A standard technique to calculate the volume of a mesh is by subdividing the shape into tetrahedrons, which are the surface triangles connected to one central point $\mathbf{x}_c$. The oriented volume of a tetrahedron is calculated by

$$V(S) = \frac{1}{6} \det\left(\mathbf{x}_1 - \mathbf{x}_c, \mathbf{x}_2 - \mathbf{x}_c, \mathbf{x}_3 - \mathbf{x}_c\right) \quad , \tag{2}$$

where the ordering of the coordinates $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ of surface triangle S determines the sign of the oriented volume. By keeping the ordering of the coordinates consistent with the normal of the surface triangles, summing all oriented volumes provides the total volume (and is independent of $\mathbf{x}_c$). However, this formulation fails when self intersections of the surface occur, since then inside-out turned parts of the shape will count as negative volume. To circumvent parts from turning inside-out during optimization, the penalty term is formulated as:

$$\mathcal{P} = \sum_{k=1}^{K} |V(S_k)| \quad , \tag{3}$$

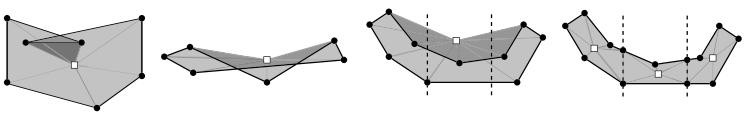


Fig. 2. Illustration of 2D self intersections and the subdivision of a non radially convex shape. White squares are centroids; darker shades of gray indicate the surfaces (i.e. volumes in 3D) that are counted multiple times in the penalty term.

where K is the number of triangles in the surface mesh. The central point $\mathbf{x}_c$ is chosen as the centroid, i.e. the average coordinate of all vertices and is recalculated during optimization. The penalty term is not a true total volume, because multiple volume parts will be counted twice or more when intersection occurs, see Figure 2. Consequently, self intersections will be penalized, since they yield a higher cost. However, shapes that are not radially convex around $\mathbf{x}_c$ will also be penalized. By this penalty shapes are forced into radial convexity. This does not need to be a restriction in practice necessarily, but it prevents structures from collapsing into a curved center line, for instance. To accomodate for this in elongated structures, a subdivision into multiple segments can be made in advance, sharing their vertices at the cut, as illustrated in Figure 2. Since the manipulation of vertex coordinates is controlled by one transformation model, the penalties can be calculated and added to the cost function per segment independently.

The gradient of the penalty term is defined

$$\frac{\partial \mathcal{P}}{\partial \mu} = \sum_{k=1}^{K} \left[\text{sign}(V(T_\mu(S_k))) \frac{\partial V(T_\mu(S_k))}{\partial \mu} \right] , \tag{4}$$

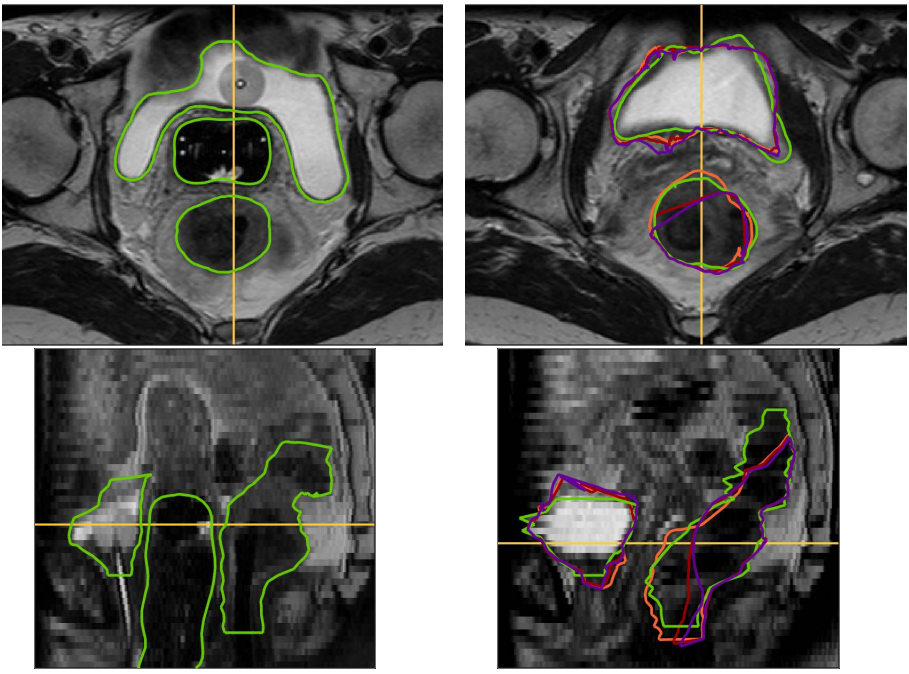
where $T_\mu(S)$ denotes $\{T_\mu(\mathbf{x}_1), T_\mu(\mathbf{x}_2), T_\mu(\mathbf{x}_3)\}$. The partial derivative $\frac{\partial}{\partial \mu} V(T_\mu(S))$ is found by applying the chain rule on the determinant, where $\frac{\partial}{\partial \mu} T_\mu(\mathbf{x})$ is a partial derivative of the deformation field. Each partial derivative to the centroid will be very small and therefore it is assumed that $\frac{\partial}{\partial \mu} \mathbf{x}_c := 0$.

3 Experiments and Results

To evaluate the effect of the proposed penalty term, five pairs of images were used. Each pair consists of one image before a brachytherapy fraction, with the applicator in situ, and one image immediately after the removal of the applicator. The scans of each pair were made about one hour apart. The images were acquired with a 1.5 T MR scanner (Gyroscan NT Intera; Philips Medical Systems, Best, The Netherlands). The image dimensions are 512×512×50 voxels, with a voxel size of 0.625×0.625×3.0 mm. Manual delineations of bladder and rectum have been provided for these images and serve as gold standard for evaluation. The bladder and rectum are organs at risk since they are close to the target volume and they are subjected to deformations caused by the applicator. The manual delineation of the applicator is used in the proposed penalty term and as a mask. Fig. 3 shows an example image with delineations.

The delineations are drawn as contours on the transversal slices. For the applicator the main body is delineated only, leaving out the needles and the intra-uterine tube (tandem) since they are small.

Registration with the proposed penalty term is compared with registration without this term. For comparison, three types of registration are performed, ordered in increasing use of prior knowledge:



(a) With applicator (b) Without applicator

Fig. 3. Example image pair (Patient 4). In image (a) along the indicated line: the bladder, the applicator and the rectum. Green delineations denote the gold standard. The propagated delineations of the experiments CnM, CwM and MSP (as described in Section 3) are shown as an overlay in purple, red and orange.

- conventional registration without applicator mask (CnM)
- conventional registration with applicator mask (CwM)
- the proposed Missing Structure Penalty (MSP)

In this study we choose normalized cross correlation as a similarity measure, because of the non-quantitative nature of the T2 images. All types of registration share the same parameters and use $R = 4$ resolution levels. The B-spline grid spacing is $\{40, 20, 10, 5\text{mm}\}$, isotropically. Since the image data are highly anisotropic, the image scale space resolutions are chosen $\{\sigma_1, \ldots, \sigma_r\} = \{16, 8, 4, 2\text{voxel}\}$ for the transversal plane and $\{4, 2, 1, 1\text{voxel}\}$ along the inferior-superior-axis. The masks that are used for experiments CwM and MSP are obtained from the applicator delineations. For experiment MSP the applicator shape is subdivided the into three parts of equal height, leading to three meshes with a total number of 229 vertices on average. The weighting factor of the penalty term is $\alpha = 10^{-8}$, this brings the penalty of unit mm^3 into the range of normalized cross correlation ($[-1, 1]$).

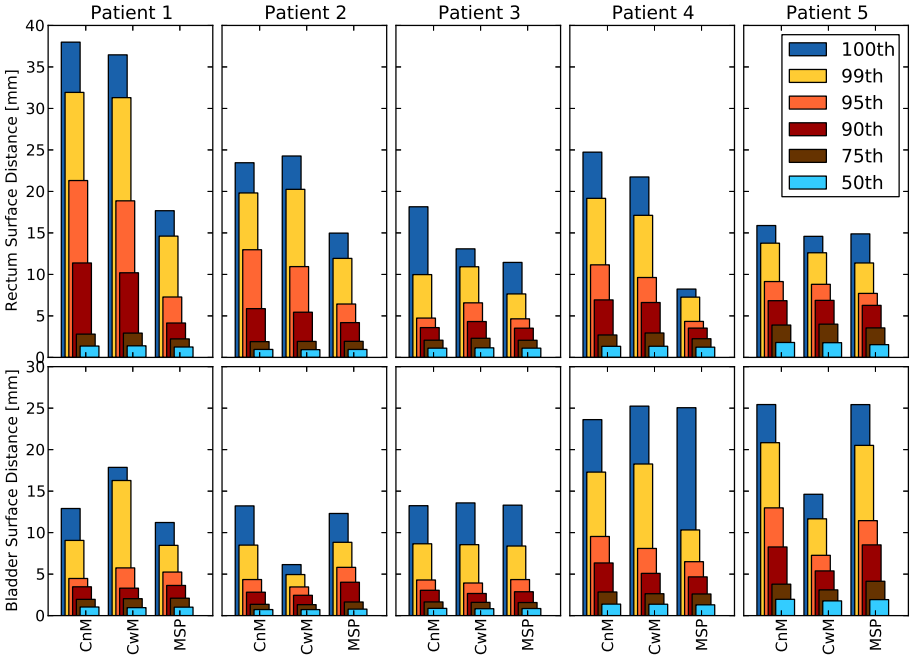
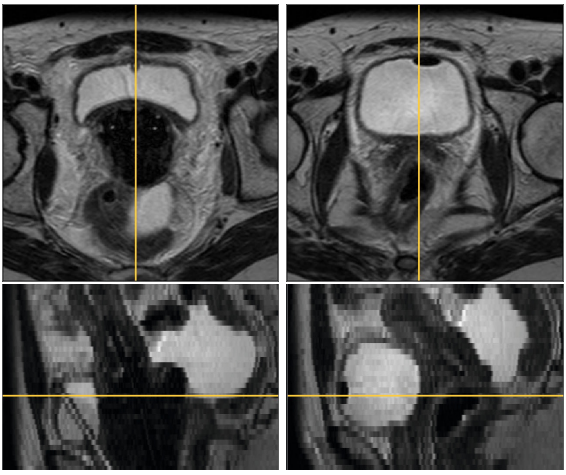


Fig. 4. Surface distance distribution percentile plots. Notice that the percentiles are not selected equidistantly.

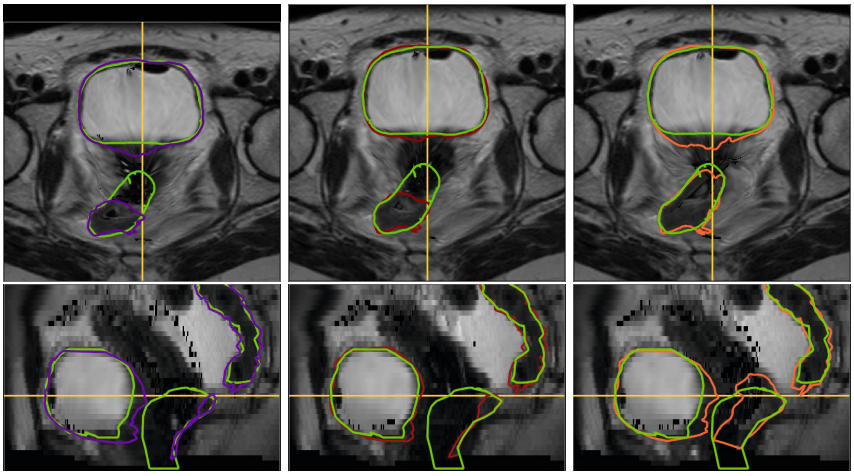
Table 1. Overlap scores and residual volumes

Experiment	Dice Rectum			Dice Bladder			Residual [ml]		
	CnM	CwM	MSP	CnM	CwM	MSP	CnM	CwM	MSP
Patient 1	0.57	0.61	0.91	0.91	0.91	0.91	51	68	2
Patient 2	0.69	0.70	0.82	0.94	0.95	0.93	29	44	5
Patient 3	0.91	0.86	0.91	0.91	0.91	0.91	23	29	2
Patient 4	0.76	0.81	0.90	0.80	0.82	0.83	26	44	1
Patient 5	0.81	0.82	0.85	0.80	0.85	0.80	12	30	3

To evaluate registration accuracy the mapping established by each experiment is used to transform the manual delineations of the rectum and bladder of the fixed image to the moving image, Figure 3b shows the delineations of one patient. A transformed delineation p^{reg} is then compared with the gold standard p^{gold}. A distribution of surface distances from p^{gold} to p^{reg} is obtained by measuring the distances of the closest plane or point on p^{reg} for all points on p^{gold} [10]. Similarly, a distribution is obtained for p^{reg} to p^{gold}. These distributions are merged into one overall distribution and are represented by bar plots indicating the percentiles of the distribution, see Figure 4. In these plots the height of the 100th percentile equals the (symmetric) Hausdorff distance.



(a) With applicator (b) Without applicator



(c) CnM (d) CwM (e) MSP

Fig. 5. Example registration results (Patient 2) by forward mapping of the image with applicator to the image without applicator. Green delineations denote the gold standard. The propagated delineations of the experiments CnM, CwM and MSP are shown as an overlay in purple, red and orange.

To avoid errors due to the varying extent in head-feet direction of the delineations of the rectum, evaluation of the surface distance is done within the range of slices where the surface meshes of all experiments were present.

A second measure of evaluation is the Dice similarity score, which indicates the amount of volume overlap. These scores are reported in Table 1, which also states the residual volume. The residual volume is obtained by voxel counting the volume of the applicator after transformation. This measure is not identical to the penalty term, but it is a true volume where negative volume due to self intersection contributes positively. Similar to the penalty term, this volume is expected to be very small for good registration.

Fig. 5 shows the registered images of the different experiments of patient 2 as an illustration of the results. The fixed image with applicator is transformed to the moving image without applicator by a forward mapping of all voxels except for the applicator region. Subsequently, the transformed voxels are linearly interpolated on the moving image grid. The residual volume of the applicator can be seen as a black line. Other black voxels in the image indicate foldings.

All registrations were done on a quad-core desktop computer running at 2.83GHz. Computation times for CnM and CwM were (both) 10 minutes and 45 seconds on average. With inclusion of the penalty term the registration time increased to 11 minutes and 12 seconds on average.

4 Discussion and Conclusion

We have proposed a method for free-form registration of two images where some structure is present in one of them, but absent in the other. We employ a penalty term for registration that minimizes the volume of the void, as for instance left by the removal of an applicator in brachytherapy of the cervix. The bladder and rectum are neighboring structures to the cervix that are affected most by this removal. Compared with conventional registration, either using a mask or not, the use of the penalty term improves the surface distance of the rectum to the gold standard. This improvement can also be seen in the Dice similarity scores. No improvement of the bladder surface distances can be seen and all methods achieved high Dice scores.

The ability of the penalty term to impose a minimal applicator void volume can best be seen by the very low residual volumes that were obtained relative to both conventional registrations. Without imposing the penalty term, the registration process could not find a mapping that has minimal applicator void volume.

The amount of folding in (and smoothness of) the deformation field is often used as a quality measure of registration. Although the total volume of folded voxels outside the applicator region is not larger when adding the proposed penalty term, folding does occur. To maintain smoothness outside of the disappearing structure but allow folding inside it, a localized bending regularization could be included in future. However, due to the applicator the uterus is forced into an upright position, while after removal the uterus returns to its tilted position. Since the uterus and the surrounding organs have sliding interfaces, this

movement induces topology changes that break the continuity and smoothness assumption of the deformation field, leading to foldings. This type of folding cannot be addressed adequately by a bending regularization and therefore remains a challenge in registration.

The outcomes of surface distances, volume overlaps and residual volume support the expectation that our approach is a proof of principle that registration of image pairs that have a structure present in one image but absent in the other, may be improved by introducing a penalty term that minimizes the void.

Acknowledgments. The authors would like to thank Christel Nomden from the Department of Radiotherapy, University Medical Center Utrecht for providing the manual delineations.

References

1. Nithiananthan, S., Schafer, S., Mirota, D.J., Stayman, J.W., Zbijewski, W., Reh, D.D., Gallia, G.L., Siewerdsen, J.H.: Extra-dimensional Demons: a method for incorporating missing tissue in deformable image registration. Med. Phys. 39(9), 5718–5731 (2012)
2. Periaswamy, S., Farid, H.: Medical image registration with partial data. Med. Image Anal. 10(3), 452–464 (2006)
3. Gu, Z., Qin, B.: Nonrigid Registration of Brain Tumor Resection MR Images Based on Joint Saliency Map and Keypoint Clustering. Sensors 9(12), 10270–10290 (2009)
4. Risholm, P., Samset, E., Talos, I.-F., Wells, W.: A non-rigid registration framework that accommodates resection and retraction. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) IPMI 2009. LNCS, vol. 5636, pp. 447–458. Springer, Heidelberg (2009)
5. Chitphakdithai, N., Duncan, J.S.: Non-rigid registration with missing correspondences in preoperative and postresection brain images. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 367–374. Springer, Heidelberg (2010)
6. Zacharaki, E.I., Hogea, C.S., Shen, D., Biros, G., Davatzikos, C.: Non-diffeomorphic registration of brain tumor images by simulating tissue loss and tumor growth. Neuroimage 46(3), 762–774 (2009)
7. Miga, M.I., Roberts, D.W., Kennedy, F.E., Platenik, L.A., Hartov, A., Lunn, K.E., Paulsen, K.D.: Modeling of retraction and resection for intraoperative updating of images. Neurosurgery 49(1), 75–84 (2001)
8. Vigneron, L.M., Noels, L., Warfield, S.K., Verly, J.G., Robe, P.A.: Serial FEM/XFEM-Based Update of Preoperative Brain Images Using Intraoperative MRI. Int. J. Biomed. Imaging 2012, 872783 (2012)
9. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging 29(1), 196–205 (2010)
10. Commandeur, F., Velut, J., Acosta, O.: A VTK algorithm for the computation of the Hausdorff distance. Midas Journal (September 2011)

Contour-Based TVUS-MR Image Registration for Mapping Small Endometrial Implants

Amir Yavariabdi[1], Chafik Samir[1], Adrien Bartoli[1], David Da Ines[1], and Nicolas Bourdel[1]

ISIT UMR 6284 CNRS/Université d'Auvergne, Clermont-Ferrand, France
{amir.yavariabdi,chafik.samir,adrien.bartoli}@gmail.com

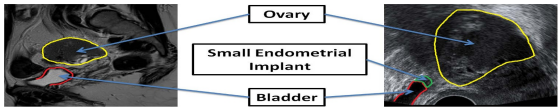
Abstract. We propose a multi-modal image registration and fusion method to cope with the limitations of Magnetic Resonance (MR) and Transvaginal Ultrasound (TVUS) imaging in observing abdominal endometrial implants. Our method facilitates the transfer of two types of information from a 2D TVUS image to a 2D MR slice: (1) the location and shape of small implants and (2) the implants' depth of infiltration in the host tissue. Our registration method uses contour correspondences through a novel variational one-step deformable Iterative Closest Point (ICP) method. The proposed method compared favorably with classical ICP and Thin-Plate Spline Robust Point Matching (TPS-RPM) on several datasets.

Keywords: Contour-based image registration, fusion, endometriosis, localization.

1 Introduction

Endometriosis is a gynecologic disorder where cells from the lining of the uterus appear and grow outside of its cavity. It is a progressive disease affecting approximately 10% of women of reproductive age. The pelvic peritoneum and pelvic organs are typical locations for endometriosis [6]. It may cause chronic pelvic pain, abnormal bleeding, dysmenorrhea, and infertility. The diagnosis of endometriosis is difficult because patients may present with a variety of symptoms. Despite of its difficulties, an accurate diagnosis should not just indicate whether a woman has endometriosis, but should also include such information as depth of infiltration and exact location. This information is necessary to plan surgery: completeness of excision highly depends on the precision of diagnosis. In other words, inaccurate information about depth of infiltration and exact location of endometrial implant may cause under- or over-cutting of the implant, the former potentially leading to recurrence. Preoperative TVUS and MR scanning combined with laparoscopic biopsy have formed the most reliable diagnosis procedure [6]. In laparoscopy diagnosis is confirmed intraoperatively. However, it does not allow the surgeon to see the depth of infiltration and implants that are located inside or under the organs. Consequently, it is important to construct an

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 145–154, 2013.



(a) MR (one slice) (b) TVUS

Fig. 1. MR imaging is limited in cases where endometriosis appears as small implants. The yellow and red curves belong to the ovary and the bladder, respectively. The green curve on TVUS shows a small endometrial implant.

endometrial map preoperatively using MR and TVUS. These imaging modalities are complementary. Small implants and depth of infiltration are only visible in TVUS while MR gives an accurate location of large implants [7] (see Figure 1). There is currently no system available to construct an endometrial map by combining MR and TVUS.

The registration and fusion between TVUS-MR data can be used to solve the limitations of MR and TVUS at observing endometriosis. This is a difficult technical problem due to the disparity in gray-level intensities, soft tissue deformations, limited field of view and low signal to noise ratio of TVUS images. Our solution to this problem takes advantage of the organs segmented by the radiologist whilst they inspect the images.

We propose a registration method that uses contour correspondences through a novel variational one-step deformable ICP method. Our registration method has two main steps: first, the MR and TVUS data are manually segmented by an expert. Second, our deformable ICP method is used to compute a dense deformation field while establishing point correspondences automatically. The novelty of this work is methodological and technical, from respectively the registration based approach to the diagnosis of endometriosis and a variational framework leading to a well-defined one-step formulation of ICP handling multiple curve correspondences.

2 Previous Works

2.1 Feature-Based US-MR Registration

Image registration has been an active topic of research for over a decade. A general survey on image registration methods is outside the scope of this paper. We refer the reader to the standard image registration textbooks by Modersitzki [2,3] or to the article by Sotiras et al. [4] for a recent comprehensive review.

US-MR feature-based registration methods can be roughly divided into three categories. (1) the methods that include automatic feature detection and matching. For example, Mitra et al. [9] register 2D Trans-Rectal US to 2D MR to assist prostate biopsy. They automatically segment the prostate contour and use a triangulation approach to generate point correspondences. The drawback of this method is that it cannot handle concave shapes. (2) the methods

requiring manual feature detection but including automatic matching. For instance, Cosse et al. [10] propose a method to register MR-US prostate and rectum surfaces. The MR image is first segmented using graph-cut while the US image is manually segmented. Registration is then carry out in two steps. First, rigid ICP is performed using the surface of the rectum and the prostate. Finally, a deformable demons algorithm is used. (3) the methods which are fully manual. For example, Reynier et al. [11] propose a feature-based registration method for brachytherapy. They manually segment contours and points in the MR and US images. Then, they elastically align the MR with the US image.

The second category seems to be a proper choice for TVUS-MR image registration. The main reason for discarding the first category is because of the tissues' characteristics at the implants. The MR image contains the full shape of pelvic organs but the corresponding organs in the TVUS images might not be fully visible. Therefore, the matching procedure would result in misplaced correspondences and the registration process would fail in estimating a precise deformation field.

2.2 ICP

Classical ICP has two main steps that are iterated until convergence: (1) closest point computation and (2) transformation estimation. ICP variant methods can be differentiated by the *type of transformation* they estimate and by the *procedure used to obtain these transformations*. As rigid ICP, deformable ICP has two inner steps. Step (2) becomes deformation estimation [5]. This can be done by minimizing an energy including a data term and a smoothing term. In order to improve the robustness and the computational efficiency of classical two-step ICP, a distance transform is introduced by Fitzgibbon et al. [12]. The distance transform allows one to merge the two inner steps. Fitzgibbon et al.'s one-step ICP computes rigid 3D/3D registration of a single pair of surfaces. The registration error is minimized using the Levenberg-Marquardt algorithm. The literature on ICP is wide and we refer the reader to Castellani et al. [1] for more details. In contrast to Fitzgibbon et al.'s one-step ICP method, ours uses a *variational procedure* to obtain *local nonlinear deformations* and handles *multiple curves correspondences*.

3 Proposed Method

The soft tissue organs and nodules in the TVUS moving image M and in the corresponding 2D MR reference slice R are manually segmented by an expert. This step constructs two sets of curves, C_M and C_R, representing the boundary of the corresponding organs in both modalities. Let Ω be a bounded open set of $\mathbb{R}^2$, $q \in C_M$ (i.e. q is a point on one of the moving curves), and $\phi \in C^2(\Omega, \mathbb{R}^2)$ the deformation vector field.

3.1 A Variational Formulation for Classical ICP

1. For every q in C_M, compute the closest-point $\zeta(q)$ on C_R:

$$\zeta(q) = \underset{p \in C_R}{\operatorname{argmin}}\, d^2(p, \phi(q)) \tag{1}$$

where ζ is a continuous function from C_M to C_R.

2. Estimate the optimal deformation vector field ϕ as:

$$\phi = \underset{\phi \in C^2(\Omega, \mathbb{R}^2)}{\operatorname{argmin}}\ \lambda \int_{C_M} d^2(\phi, \zeta)\, dq + (1-\lambda) \int_{\Omega} \|\triangle\phi\|_2^2\, dX \tag{2}$$

where $\triangle$ is the Laplacian operator and $\lambda \in [0,1]$ is a smoothing parameter. The two terms in equation (2) are the data term and the smoothing term.

3.2 Variational Formulation Using Distance Transform

Substituting equation (1) in (2) yields:

$$\phi = \underset{\phi \in C^2(\Omega, \mathbb{R}^2)}{\operatorname{argmin}}\ \lambda \int_{C_M} d^2\left(\phi, \underset{p \in C_R}{\operatorname{argmin}}\ d^2(p, \phi)\right) dq + (1-\lambda) \int_{\Omega} \|\triangle\phi\|_2^2\, dX \tag{3}$$

It is clear that $d^2\left(\phi, \underset{p \in C_R}{\operatorname{argmin}}\ d^2(p, \phi)\right) = \min\limits_{p \in C_R} d^2(p, \phi)$. This can be interpreted as the fact that the cost depends on the distance to the closest point but not on the closest point itself. This converts the problem into:

$$\phi = \underset{\phi \in C^2(\Omega, \mathbb{R}^2)}{\operatorname{argmin}}\ \lambda \int_{C_M} \min_{p \in C_R} d^2(p, \phi)\, dq + (1-\lambda) \int_{\Omega} \|\triangle\phi\|_2^2\, dX \tag{4}$$

The data term now involves a Distance Transform, $D \circ \phi = \min\limits_{p \in C_R} d^2(p, \phi)$. This results in:

$$\phi = \underset{\phi \in C^2(\Omega, \mathbb{R}^2)}{\operatorname{argmin}}\ \lambda \int_{C_M} (D \circ \phi)^2\, dq + (1-\lambda) \int_{\Omega} \|\triangle\phi\|_2^2\, dX \tag{5}$$

3.3 Numerical Solution

To solve the nonlinear variational problem (5), several numerical optimization algorithms can be applied [2,3]. In equation (5), the data term places constraints on the deformation field ϕ at the contour locations. To make it more general and to permit its formulation as a Euler-Lagrange (EL) equation we use an index function δ_{C_M}, with $\delta_{C_M} : \Omega \to \{0,1\}$, $\delta_{C_M}(X) = 1$ if $X \in C_M$ and 0 otherwise. We rewrite the cost functional of equation (5) as:

$$E[\phi] = \int_{\Omega} \left(\lambda \delta_{C_M} (D \circ \phi)^2 + (1-\lambda) \|\triangle\phi\|_2^2\right) dX \tag{6}$$

Let ϕ^0 be an initial estimate that can be found by rigid registration. We follow the iteration:

$$\phi^{k+1} = \phi^k + U \tag{7}$$

where $U = [u_1, u_2]^T$. By substituting equation (7) into equation (6), we obtain:

$$E[\phi^k + U] = \int_{\Omega} \left(\lambda \delta_{C_M} \left(D \circ \left(\phi^k + U \right) \right)^2 + \left((1-\lambda) \left\| \triangle U \right\|_2^2 + \left\| \triangle \phi^k \right\|_2^2 \right) \right) dX \tag{8}$$

The distance transform D is nonlinear and can be approximated by its first order Taylor expansion.

$$E[\phi^k + U] = \int_{\Omega} \left(\lambda \delta_{C_M} \left(D \circ \phi^k + \left(\nabla D \circ \phi^k \right) U \right)^2 + \left((1-\lambda) \left\| \triangle U \right\|_2^2 + \left\| \triangle \phi^k \right\|_2^2 \right) \right) dX \tag{9}$$

where $\nabla = [\frac{\partial}{\partial x}, \frac{\partial}{\partial y}]$. Using EL, a system of two fourth order elliptic Partial Differential Equations (PDEs) is derived:

$$\mu \int_{C_M} \left(D \circ \phi^k + \left(\nabla D \circ \phi^k \right) U \right) \left(\nabla D \circ \phi^k \right) dq + \left(\frac{\partial^4 U}{\partial X^4} \right) = 0 \tag{10}$$

where $\mu = \frac{\lambda}{1-\lambda}$. We discretize the curves C_M in N points $q_1, \ldots, q_N$ and Ω on the pixel grid. Note that $\delta_q(X) = 1$ if $X = q_j$ for $j = 1, \ldots, N$ and 0 otherwise.

$$\mu \delta_q \left(X \right) \Bigg(\left(D \left(\phi^k \left(X \right) \right) + D_x \left(\phi^k \left(X \right) \right) u_1 \left(X \right) + D_y \left(\phi^k \left(X \right) \right) u_2 \left(X \right) \right) D_\ell \left(\phi^k \left(X \right) \right) \Bigg) + \left(\frac{\partial^4 U}{\partial x^4} + \frac{\partial^4 U}{\partial y^4} \right) = 0 \tag{11}$$

where D_ℓ is the derivative of D with respect to $\ell \in \{x, y\}$. To solve the PDEs, we use a finite difference scheme with boundary condition $\triangle U = 0$. This leads to a sparse linear system that can be solved by Successive Over-Relaxation (SOR):

$$\begin{aligned} u_{1,h,w}^{t+1} = \beta u_{1,h,w}^t - A \Bigg[& \mu \delta_q \left(X \right) \left(D \left(\phi^k \left(X \right) \right) \frac{\partial D}{\partial x} \left(\phi^k \left(X \right) \right) \right) + C u_{2,h,w}^t \\ & + \left(u_{1,h,w-2}^{t+1} - 4 u_{1,h,w-1}^{t+1} - 4 u_{1,h+1,w}^{t+1} + u_{1,h+2,w}^{t+1} \right) \\ & + \left(u_{1,h,w+2}^t - 4 u_{1,h,w+1}^t - 4 u_{1,h-1,w}^t + u_{1,h-2,w}^t \right) \Bigg] \\ u_{2,h,w}^{t+1} = \beta u_{2,h,w}^t - B \Bigg[& \mu \delta_q \left(X \right) \left(D \left(\phi^k \left(X \right) \right) \frac{\partial D}{\partial y} \left(\phi^k \left(X \right) \right) \right) + C u_{1,h,w}^t \\ & + \left(u_{2,h,w-2}^{t+1} - 4 u_{2,h,w-1}^{t+1} - 4 u_{2,h+1,w}^{t+1} + u_{2,h+2,w}^{t+1} \right) \\ & + \left(u_{2,h,w+2}^t - 4 u_{2,h,w+1}^t - 4 u_{2,h-1,w}^t + u_{2,h-2,w}^t \right) \Bigg] \end{aligned} \tag{12}$$

where $A = \frac{\alpha}{12+\mu\delta_q(X)\left(\frac{\partial D}{\partial x}\left(\phi^k(X)\right)\right)^2}$, $B = \frac{\alpha}{12+\mu\delta_q(X)\left(\frac{\partial D}{\partial y}\left(\phi^k(X)\right)\right)^2}$, $t \in \mathbb{N}$ is the iteration number, $C = \mu\delta_q(X)\left(\frac{\partial D}{\partial y}\left(\phi^k(X)\right)\frac{\partial D}{\partial x}\left(\phi^k(X)\right)\right)$, $\beta = 1-\alpha$, $h = 1,\ldots,H$, $w = 1,\ldots,W$, H and W are height and width of M respectively, and α is the relaxation factor.

4 Experimental Results

The 2D MR slice from the MR volume which corresponds to the TVUS slice is manually selected by an expert. We evaluate the proposed method's performance in comparison with variational two-step deformable ICP, TPS-RPM [13] which is one of the best non-rigid point-set registration methods, and a point-based registration method [8]. We illustrate the registration results on various semi-synthetic and real-world data. Registration accuracy is evaluated in terms of Dice Similarity Coefficient (DSC), Hausdorff Distance (HD), Target Registration Error (TRE), Correlation Coefficient (CC), and Mean Square Error (MSE).

DSC assesses the global overlap of the segmented organs. A high DSC value indicates a good overlap between the tissue regions after registration. However, a high DSC does not mean a good contour overlap. Therefore, HD is used to evaluate the contour accuracy. A low HD value shows good contour overlap. TRE measures the distance after registration between corresponding points not used in estimating the deformation. Notably a low TRE value shows good local registration accuracy. In our experiments, the target points are 8 boundary points which are not used for registration.

Numerous experiments on semi-synthetic data with various degrees of deformation are performed to evaluate one-step ICP's performance. We evaluate the proposed method against two-step ICP and TPS-RPM. In order to generate our semi-synthetic data, the boundary of soft tissue organs such as the bladder, the uterus, the ovary, and the rectum are manually segmented from different MR slices. Then, randomly generated deformations are applied to each curve, and different algorithms are used to recover those deformations. Note that the contours are deformed through increasingly larger degrees of deformation (see Figure 2, first row). The rate of deformation varies between 15% and 75%. The MSE is used to estimate the error between corresponding points. 20 2D contours with 5 degrees of deformation are used to evaluate the accuracy of the methods. Both one-step ICP and TPS-RPM provide accurate results. Two-step ICP gets trapped in a local minimum. As shown in Figure 2 (e), one-step ICP shows robustness even in areas of complex deformation. Figure 3 shows one-step ICP, TPS-RPM, and two-step ICP performances for varying levels of deformation. In Figure 3, error means and standard deviations for each setting are shown. As the closest point pairs are estimated through the registration process, we use them to evaluate the accuracy of fitting (Figure 3 (a)). In addition since the true correspondences are known, we can estimate the registration error (see Figure 3 (b)). One-step ICP outperforms TPS-RPM. Two-step ICP fails in most of the experiments. Based on these experiments, we conclude that our method has high registration accuracy and outperforms TPS-RPM.

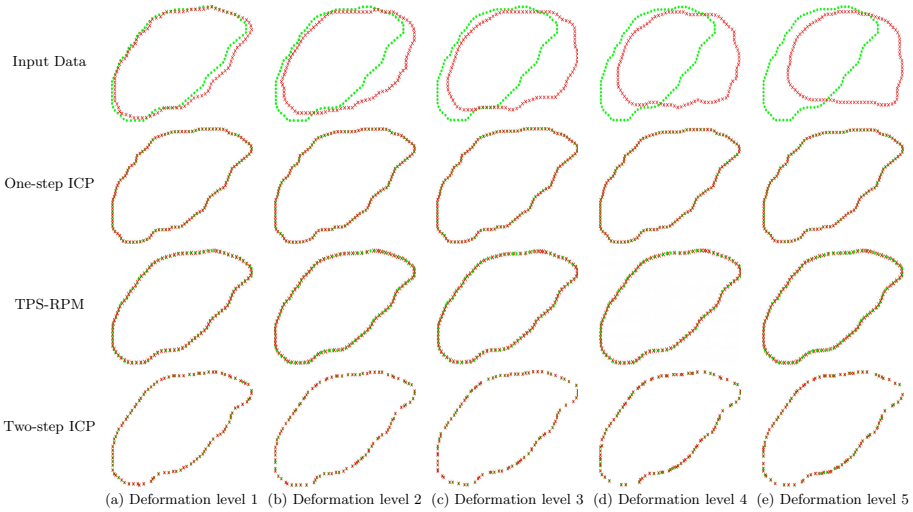
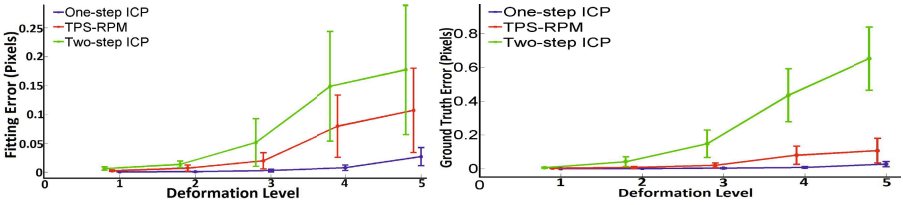


Fig. 2. Registration results for semi-synthetic contours. The proposed method (second row) is compared with two-step ICP (last row) and TPS-RPM (third row). The first row shows the reference (•) and moving (×) point sets. Each column shows a deformation level. Best viewed in colour and close-up.



(a) MSE between the closest point pairs (b) MSE between the true correspondences

Fig. 3. Comparison of one-step ICP with different methods on the semi-synthetic data with respect to deformation level. Our method achieves high registration accuracy compared to the two other methods.

For real-world data, we use a point-based registration method [8] as a baseline to evaluate our results. In this manner, an expert manually selects control points at each pair of curves. The reasons that we chose [8] is a comparison with the well-known method of the state-of-the-art are given therein. The algorithms are implemented in Matlab and tested on an i5 core 3.3 GHz with 16 GB RAM. The registration accuracy that measures overlap are tabulated in Table 1. It shows that our method outperforms the other methods in terms of contour overlap accuracy. TRE and computational time are tabulated in Table 2. It is indicated that the proposed method provides high local registration accuracy and is fast. Tables 1 and 2 indicate that the proposed method provides promising results as the difference with the baseline method is the smallest one compared to other

methods. Moreover, two experts compared our method with the baseline method. In this manner, we setup a registration quality score called 1-to-5 rating scale where 5 is excellent and 1 is bad. The experts' evaluations are given in Table 3.

Figures 4 and 5 show the registration results for two different patients. The registration results for the point-based registration method [8], Proposed method, TPS-RPM, and two-step ICP are illustrated in rows. The first column illustrates the reference and moving images. In Figure 4, the depth of infiltration is not clear in the MR image. In Figure 5, endometrial tissues are seen in the TVUS image and there is no evidence of their presence in the MR image. For localizing the endometriosis we apply the displacement field to the yellow curve in the TVUS images and we find the new location of endometriosis in the MR images. From

Table 1. Comparison between point-based registration, proposed method, TPS-RPM and two-step ICP. A high DSC value means a good contour region overlap, while a low HD value signifies a good contour overlap. **Bold** values indicate the best results.

	DSC				HD			
Patient	Point-based	Proposed	TPS-RPM	2-step ICP	Point-based	Proposed	TPS-RPM	2-step ICP
1	0.9943	**0.9939**	0.9934	0.9842	0.0799	**1.2632**	1.5661	4.1604
2	0.9897	**0.9893**	0.9867	0.9816	0.2873	**1.2605**	1.5345	1.9078
3	0.9946	**0.9800**	0.9771	0.9427	0.8088	**1.6623**	1.8384	3.6525
4	0.9819	**0.9626**	0.9147	0.8195	0.7231	**1.7258**	2.4013	10.8844
5	0.9960	**0.9948**	0.9910	0.9874	0.7546	**1.7821**	2.0539	2.6033
average	0.9913	**0.9841**	0.9726	0.9431	0.5307	**1.5388**	1.8788	4.6417
std. dev.	0.0058	**0.0134**	0.0329	0.0714	0.3267	**0.2563**	0.3611	3.599

Table 2. Comparison between point-based registration method (baseline), proposed method, TPS-RPM and two-step ICP. A low TRE value means good local registration accuracies around target landmarks. **Bold** values indicate the best results.

	TRE				Time(sec)			
Patient	Point-based	Proposed	TPS-RPM	2-step ICP	Point-based	Proposed	TPS-RPM	2-step ICP
1	0.17	**0.7880**	1.8895	2.2052	7.52	**12.79**	77.53	14.13
2	0.263	**0.5061**	1.4825	1.7540	16.12	**18.37**	90.64	24.34
3	1.1153	**2.2757**	4.2002	6.6531	9.59	**22.68**	83.03	27.96
4	0.1929	**1.2486**	3.1181	8.1541	7.13	**14.78**	94.82	15.69
5	0.651	**0.7289**	0.9	1.581	11.03	**13.07**	96.92	15.31
average	0.4784	**1.1095**	2.3181	4.0695	10.2780	**16.3380**	88.5880	19.4860
std. dev.	0.4057	**0.7056**	1.3300	3.0979	3.6292	**4.1847**	8.1493	6.2431

Table 3. Experts' evaluations of the point-based registration and proposed method

	Radiologist		Surgeon	
Patient	Point-based	Proposed	Point-based	Proposed
1	4	4	4	5
2	3	4	4	4
3	2	3	3	3
4	5	5	4	5
5	4	3	4	4
average	3.6	3.8	3.8	4.2
std. dev.	1.14	0.83	0.2	0.7

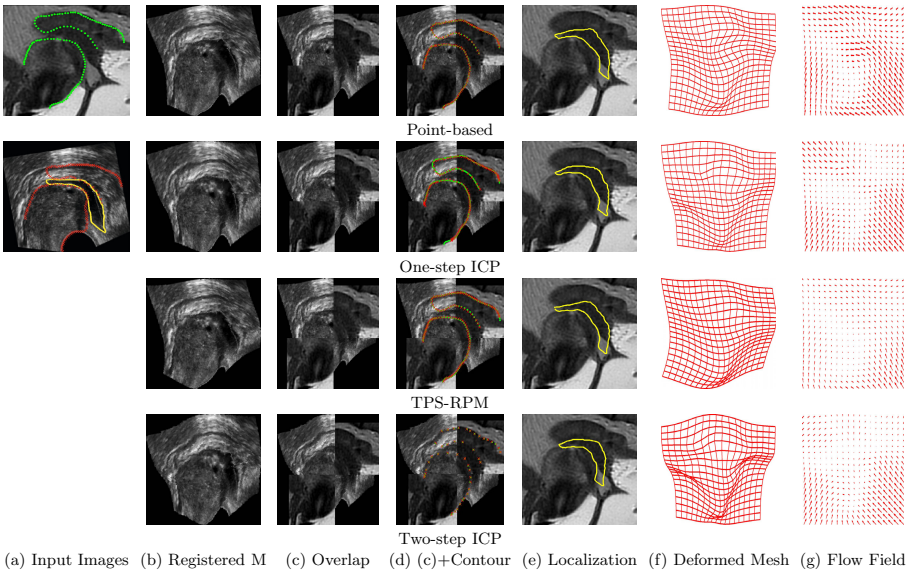


Fig. 4. Registration results for patient 1. Columns 3 and 4 show the overlap images where the top right and the bottom left belong to the MR image and the top left and the bottom right belong to the deformed TVUS image.

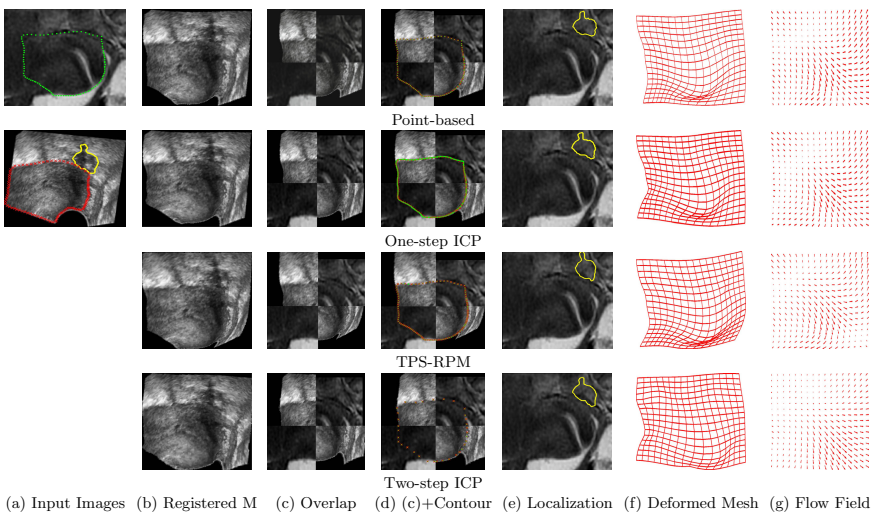


Fig. 5. Registration results for patient 2

Figures 4 and 5 we can conclude that the proposed registration method provide an accurate and smooth displacement field to localize the implants.

Finally, the evaluations and the experimental results show that the proposed method gives accurate and smooth displacement fields. Moreover, they demonstrate that the combination of complimentary information from TVUS and MR

images is more informative than any of the input modalities. Indeed, the registration provides more precision about implant location and depth of infiltration which consequently simplifies and improves diagnosis.

5 Conclusion

A new framework for 2D/2D TVUS-MR registration is proposed to localize and characterize endometrial tissues. The method uses contour correspondences through a novel variational one-step ICP. In order to obtain point correspondences and local nonlinear deformations, the method uses Euclidean distance maps resulting from MR contours. The performance of the proposed method is evaluated against variational two-step deformable ICP and TPS-RPM. The results obtained from semi-synthetic and real-world data conclude that the performance of the proposed method is better than the two other methods.

References

1. Castellani, U., Bartoli, A.: 3D Shape Registration. In: 3D Imaging, Analysis, and Applications (2012)
2. Modersitzki, J.: Numerical Methods for Image Registration. Oxford University Press (2004)
3. Modersitzki, J.: Flexible Algorithms for Image Registration. SIAM (2009)
4. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable Medical Image Registration: A Survey. In: TMI, pp. 1153–1190. IEEE Press, New York (2013)
5. Amberg, B., Romdhani, S., Vetter, T.: Optimal Step Nonrigid ICP Algorithm for Surface Registration. In: CVPR (2007)
6. Chamie, L.P., Blasbalg, R., Pereira, R.M.A., Warmbrand, G., Serafini, P.C.: Findings of Pelvic Endometriosis at Transvaginal US, MR Imaging, and Laproscopy. RadioGraphics 31, 71–100 (2011)
7. Brosens, I., Puttemans, P., Campo, R., Gordts, S., Kinkel, K.: Diagnosis of Endometriosis: Pelvic Endoscopy and Imaging Techniques. Best Practice and Research Clinical Obstetrics and Gynaecology 18, 285–303 (2004)
8. Yavariabdi, A., Samir, C., Bartoli, A., Da Ines, D., Bourdel, N.: Mapping Endometrial Implants by 2D/2D Registration of TVUS to MR Images from Point Correspondences. In: ISBI (2013)
9. Mitra, J., Oliver, A., Marti, R., Llado, X., Vilanova, J.C.: Multimodal Prostate Registration using Thin-Plate Splines from Automatic Correspondences. Digital Image Computing: Techniques and Applications (2010)
10. Cosse, A.: Diffeomorphic Surface-based Registration for MR-US Fusion in Prostate Brachytherapy. In: MELECON, pp. 903–907 (2012)
11. Reynier, C., Troccaz, J., Fourneret, P., Dusserre, A., Gayjeune, C., Descotes, J., Bolla, M., Giraud, J.: MRI/TRUS data fusion for prostate brachytherapy. Preliminary results. Medical Physics 31, 1568–1575 (2004)
12. Fitzgibbon, A.W., Levoy, M.: Robust Registration of 2D and 3D Point Sets. Image and Vision Computing, 1145–1153 (2003)
13. Chui, H., Rangarajan, A.: A New Point Matching Algorithm for Non-Rigid Registration. CVIU 89, 114–141 (2003)

Rigid Registration of Untracked Freehand 2D Ultrasound Sweeps to 3D CT of Liver Tumours

Amalia Cifor[1], Laurent Risser[2], Mattias P. Heinrich[1], Daniel Chung[1,3], and Julia A. Schnabel[1]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Old Road Campus Research Building, OX3 7DQ, Oxford, UK
amalia.cifor@eng.ox.ac.uk
[2] CNRS, Toulouse Mathematics Institute, Paul Sabatier University, 118 Route de Narbonne, 31400, Toulouse, France
[3] Department of Radiology, Churchill Hospital, Old Road, OX3 7LE, Oxford, UK

Abstract. We present a rigid registration framework for freehand 2D ultrasound sweeps to 3D CT of liver tumours. The method registers the 2D sweeps in a group-wise manner, without the need for prior 3D ultrasound compounding or probe tracking during acquisition. We first introduce a specific acquisition model to keep the dimension of this problem reasonable. Only seven parameters are indeed required to register the images. These are estimated using simulated annealing optimization of a robust modality-independent similarity measure. The framework contrasts the current methods that rely on tracking devices and phantom calibration, which are often difficult to use routinely in clinical practice. Our results on both synthetic and real data show that the method is well-suited for such ultrasound-CT registration of liver tumours.

Keywords: registration, multimodal, ultrasound, computed tomography, liver tumour.

1 Introduction

Much improvement in diagnoses, surgical planning, image guided-intervention, and treatment monitoring is attributed to the development of image fusion techniques of complementary modalities. By registering images of various intrinsic characteristics, we are able to view the anatomical structures in multiple spatio-temporal dimensions (e.g. with 3D+t ultrasound (US) and 3D magnetic resonance imaging (MRI)), and simultaneously at structural and functional levels (e.g. positron emission tomography (PET) - computed tomography (CT)). We aim at combining the merits of both 2D US and 3D CT to assess the liver tumour response to treatment in patients undergoing chemotherapy. A rigid registration between the two modalities is a first requirement in our clinical study, whose motivation is to determine the efficacy of ongoing treatment. While multiple CT acquisitions would be preferable to investigate the structural changes of liver tumours over the treatment period, this is clinically infeasible due to ionising

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 155–164, 2013.

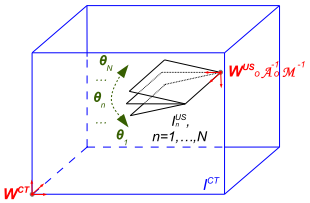


Fig. 1. Summary of the notations. A series of 2D US frames $I_n^{US}, n = 1, \cdots, N$ is registered to a 3D CT image I^{CT}. To map the images in world coordinates, we use: **(a)** I^{CT} location in the world domain. It depends on I^{CT}'s image to world matrix $\mathbf{W}^{CT}$. **(b)** $I_{n_{ref}}^{US}$ location in the world domain. It depends on $I_{n_{ref}}^{US}$'s image to world matrix $\mathbf{W}^{US}$, the manually defined transformation matrix $\boldsymbol{\mathcal{M}}$ and the *estimated* adjustment matrix $\boldsymbol{\mathcal{A}}$. **(c)** *Estimated* rotation matrices $\boldsymbol{\Theta}_n$ define the aperture of the fan.

radiation and cost factors involved. In contrast, US is much more affordable, safe and fast. By registering frequently acquired US scans with the CT volumes obtained in standard clinical practice, we offer a framework in which the tumour response to treatment could be assessed repeatedly, rather than at the end of the chemotherapy. This assessment would then facilitate a personalized treatment, in which the drug dosage is adjusted over the course of the treatment according to the tumour response.

Although 3D US probes are emergent, 2D US acquisition is still preferred in liver imaging for clinical purposes. Registration of 2D US and 3D CT is challenging due to the following: (1) the input data have different dimensionality; (2) the images have different noise models; and (3) the anatomical structures appear different in the two images (e.g. the vessels are bright in contrast enhanced CT and dark in US). The techniques tailored for such US-CT registration usually use a set of 2D US frames acquired with a freehand sweep movement of the probe, tracked with an external device. The information supplied by the tracking device is then used twofold: to build a 3D US volume from the sweep [1], and to establish 3D point correspondences between the US and CT spaces using a calibration method [2]. The multi-modal US to CT/MRI registrations then rely on robust similarity measures (e.g. adapted correlation ratio [3]), features [4], and intermediary simulated US-like images [2,5].

The position tracking devices are costly and have limited use in clinical practice. Therefore, current research efforts aim at estimating the freehand probe movement using the image data. Such techniques use speckle decorrelation to estimate the fan angle (i.e. the angle under which the sweep was acquired) from the set of frames [6], and fusion of patch-based correlation measurements [7]. These approaches however require prior phantom analysis for calibration.

In this paper we propose a rigid registration method for untracked 2D US sweeps to 3D CT, without the need for phantom calibration. We use the set of US frames in a group-wise manner, and simultaneously register them with the 3D CT, based on the following assumptions about the US sweep acquisitions: (1) the position of the probe is fixed during the sweep movement; and (2) the

sweep is performed smoothly along a certain direction. Our method consists of estimating 7 parameters: a 3D rigid body transformation for the position and orientation of the probe, and a rotation fan angle. The parameters are estimated with simulated annealing, by optimizing a robust multi-modal similarity measure: Modality Independent Neighbourhood Descriptor (MIND) [8]. Section 2 details the components of our method. Then, Section 3 shows the potential use of our method in 2D US sweeps to 3D CT registration of liver tumours, based on its performance on real and synthetic data. We conclude with an outlook of our work towards its use in clinical trials.

2 Method

The proposed framework aims at finding the rigid transformation that aligns a source 3D CT volume, I^{CT}, with the target US sweep formed of a set of N 2D frames, $I^{US} = \{I_n^{US}, n = 1, \cdots, N\}$. To solve the registration problem, we first design a specific model summarized in Fig. 1 and described in detail bellow. Registration of I^{CT} onto I_n^{US} is then performed in two steps: initialization and automatic rigid registration.

In the first step, a clinical expert performs a *rough* manual initialization of the position and orientation of the US probe using a reference US image $I_{n_{ref}}^{US}$ with respect to I^{CT}. The outcome of this initialization is a 3D rigid transformation matrix denoted here by $\boldsymbol{\mathcal{M}}$.

In the second step, we automatically adjust this manual rigid transformation and estimate the location of all US images I_n^{US} relative to $I_{n_{ref}}^{US}$, in a group-wise registration manner. In practice, this consists of automatic estimation of 7 parameters. The 3D rigid body transformation is modelled by an adjustment matrix $\boldsymbol{\mathcal{A}}$ which is encoded by 3 rotation angles ρ_x, ρ_y, ρ_z and 3 translations t_x, t_y, t_z. It aims at improving the manual estimate (matrix $\boldsymbol{\mathcal{M}}$) of the position and orientation of the probe using the entire sweep, instead of one frame only. The 7^{th} parameter is a rotation angle θ which gives the size of the fan. Since the sweep was obtained with a smooth movement in one direction, we assume that the positions of the frames in the US sweep are almost linearly sampled. Consequently, each frame position in the sweep can be identified analytically by the following equation:

$$\theta_n = \pm\theta\frac{(n - n_{ref})}{N} \tag{1}$$

where $\pm$ indicates the direction of the sweep (e.g. up-down or down-up) and n_{ref} is the index of the reference US frame used in the initialization. Using Eq. (1), we then define a set of 3D one-angled rotation matrices $\boldsymbol{\Theta}_n$, $n = 1, \cdots, N$ (see Fig. 1). To compare the images, we finally introduce the transformation $\boldsymbol{\mathcal{T}}_n$ that yields the CT plane corresponding to the n^{th} US frame in the sweep. It is defined by the composition of the three matrices:

$$\mathcal{T}_n = \boldsymbol{\Theta}_n \circ \boldsymbol{\mathcal{A}} \circ \boldsymbol{\mathcal{M}} = \begin{bmatrix} \theta_n & 0 \\ 0 & 1 \end{bmatrix} \circ \underbrace{\begin{bmatrix} R_a & T_a \\ 0 & 1 \end{bmatrix}}_{\boldsymbol{\mathcal{A}}} \circ \underbrace{\begin{bmatrix} R_m & T_m \\ 0 & 1 \end{bmatrix}}_{\boldsymbol{\mathcal{M}}} \tag{2}$$

where R_m, R_a, T_m, and T_a are 3D rotation and translation parameters which approximate the position and orientation of the probe relative to the CT image location. Before applying this transformation to the CT image, we also account for the different voxel sizes in the two modalities. We consider $\mathbf{W}^{CT}$ and $\mathbf{W}^{US}$ the transformation matrices which convert the voxel coordinates into world coordinates (in millimetres) in I^{CT} and $I_n^{US}, n = 1..N$, respectively. We can then resample I^{CT} in a 2D plane corresponding to the image domain of I_n^{US} using:

$$I_n^{CT} = I^{CT} \circ (\mathbf{W}^{CT})^{-1} \circ (\boldsymbol{\mathcal{T}}_n)^{-1} \circ \mathbf{W}^{US} \tag{3}$$

This enables us to measure the similarity between the registered 3D CT image and the N 2D US images.

2.1 Similarity Measure

Modelling the differences between modality-dependent noise distributions, contrast and structural appearance, is often difficult in the registration of US images with either US or another modality. Some recent registration methods show that structural features (e.g. in US-US registration [9]) or descriptors (e.g. in CT-MRI registration [8]) yield more robust similarity measures than the commonly used intensity values. The advantage of using such descriptors is that they offer a modality-independent representation of the image content [8], or contrast-invariant structural information [9]. In this paper, we compare the input images with MIND [8], which is a multi-modal similarity measure derived from locally-estimated structural descriptors, and originally designed for MRI-CT registration. The local structural descriptors of an image I are based on self-similarity distances of small image patches within small regions R:

$$\mathcal{D}^I(\mathbf{x}, \mathbf{r}) = \frac{1}{f} \exp\left(-\frac{\sum_{\mathbf{p}\in P}(I(\mathbf{x}+\mathbf{p}) - I(\mathbf{x}+\mathbf{r}+\mathbf{p}))^2}{var}\right) \tag{4}$$

Here, $\mathbf{r}$ is the spatial distance between a patch centred at a voxel $\mathbf{x}$ and another patch in its neighbourhood R. P is the size of the patch and defines the scale of detected features. f is a normalization factor, and var is a local noise estimate which controls the amount of local structural information captured by the descriptor. A suitable value of var is the average of all patch distances within the neighbourhood. We adopted a fast implementation of the distance measure in Eq. (4) using the convolution scheme proposed in [8].

We compute the descriptors for each 2D US frame and corresponding (2D) resampled CT plane. We choose a local neighbourhood $R = 8$, which allows us to identify features in all immediate directions from a given 2D patch. The similarity measure is then the sum of absolute differences (SAD) of the set of descriptors corresponding to each pair of frames:

$$S_n(\mathbf{x}) = \frac{1}{|R|} \sum_{\mathbf{r}\in R} |\mathcal{D}^{I_n^{US}}(\mathbf{x}, \mathbf{r}) - \mathcal{D}^{I_n^{CT}}(\mathbf{x}, \mathbf{r})| \tag{5}$$

Note that MIND also allows to use mono-modal specific similarity measures (e.g. SAD) in our multi-modal US-CT registration.

2.2 Rigid Registration with Simulated Annealing

Adopting the sweep transformation model and similarity measure described above, we then estimate the 7 parameters simultaneously using simulated annealing technique [10]. The merits of this classic optimization method can be summarized in the following: (1) it finds an approximation of the global optimum from among several local extrema; and (2) the exploration of the parameter space is independent of the optimization landscape. We note that similar philosophy is adopted in the discrete optimization registration of Zikic et al. [11]. The simulated annealing process resembles the slow cooling effect in thermodynamics. Similarly to a physical system that reaches the minimum energy state as a result of a slow decrease in temperature (e.g. crystallization), the optimization converges to the global extrema of a function after a slow decrease in a temperature parameter inherent to the simulated annealing technique. Essentially, the solution approximates the global extremum after passing many extrema during a number of iterations. The method has three key components: the *temperature* which is initialized to a high value, an *annealing schedule* which defines how the temperature is decreased, and an *acceptance probability function* which determines whether the generated solution should be accepted or not. The purpose of the acceptance function is to prohibit the optimization from being caught in a local minimum. Our method then minimizes the measure in Eq. (5) for $n = 1..N$, using Algorithm 1 below:

Algorithm 1. Group-wise rigid registration of 2D US sweeps to 3D CT

$p_{current} = \{\rho_x, \rho_y, \rho_z, t_x, t_y, t_z, \theta\} = \{0, 0, 0, 0, 0, 0, 0\}$
$S_{current} = maxValue$
$temp = temp_0$
for $i_t = 1$ to m_t **do**
  for $iteration = 1$ to m **do**
    {*Generate new test values for the 7 parameters*}
    $p_{test} = p_{current} + C(\mathbf{x})$
    Compute the similarity S_n in (5) for all corresponding frames given p_{test}
    $S_{test} = \sum_{n=1}^{N} S_n$
    if $S_{test} < S_{current}$ **then**
      {*Keep the generated values*}
      $p_{current} = p_{test}$
      $S_{current} = S_{test}$
    else
      if $\delta \leq \exp(-(S_{test} - S_{current})/(k \cdot temp)$ **then**
        $p_{current} = p_{test}$
        $S_{current} = S_{test}$
      end if
    end if
  end for
  $temp = temp_0/(i_t + 1)$
end for

We chose a fast simulated annealing implementation [12], in which the temperature *temp* is decreased inversely linear in time after m iterations (line 21), and the new parameter values are randomly generated with the Cauchy function:

$$C(x) = \frac{temp}{(x^2 + temp^2)} \tag{6}$$

where x is a randomly generated number. We use a typical Metropolis test in line 15, where the constant k controls the scale of the uphill jumps. δ is a random number obtained from a uniform distribution in $[0, 1]$, which determines the probability of acceptance of the estimated solution. $p_{current}$ denotes the set of estimated parameters, obtained by registering all frames in the sweep in a group-wise manner (line 9).

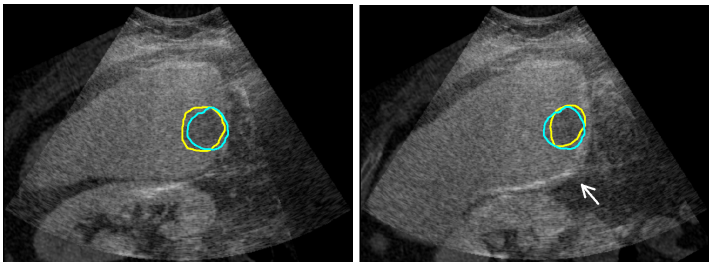
3 Experiments and Results

We tested the proposed algorithm on both synthetic and real datasets. We first show the performance of our method on simulated data, and then present the registration results on US-CT images of 5 patients. In all experiments, the constant k is 0.02, and the size of the patches in MIND is $P = [5 \times 5]$.

In the synthetic data experiments, the target image is a sweep-like set of 10 frames extracted from a CT volume with a known fan angle and added multiplicative noise, which simulates the noise distribution in real US images. The random noise is generated with a uniform distribution of zero mean and variance 0.04. The fan angle is $\theta_{target} \in \{20, 30, 40, 50, 60\}$ degrees, which produces sweeps large enough to capture the entire visible tumour and parts of the liver. These sweeps closely mimic the possible US sweep acquisitions for this particular patient. The source image is the original CT data, without additional noise. In order to simulate the errors that may be caused by the manual initialization, we initialize $\boldsymbol{\mathcal{A}}$ with a random rigid transformation whose rotation and translation parameters are obtained from a Cauchy distribution of $\sigma = 5.0$. Similarly, we give an initial fan estimate, θ_{source}, which is a random deviation from 20 degrees, also obtained from a Cauchy distribution of $\sigma = 5.0$. Note that this initialization is equivalent to perturbing the CT source volume with a random transformation.

Table 1. Mean and standard deviation of the TRE (mm) and of the perturbed source fan angle, θ_{source}, before and after registration of synthetic data; θ_{target} is the fan angle of the target sweep

θ_{target}	Mean and std of θ_{source}		Mean and std of TRE (mm)	
	before reg	after reg	before reg	after reg
20	23.6 ± 16.2	**20.0 ± 0.2**	22.2 ± 11.1	0.2 ± 0.1
30	21.1 ± 10.3	**29.9 ± 0.2**	27.6 ± 13.4	0.2 ± 0.0
40	31.1 ± 22.4	**39.9 ± 0.4**	20.4 ± 8.8	0.18 ± 0.0
50	24.4 ± 9.7	**49.7 ± 0.4**	25.8 ± 12.7	0.2 ± 0.0
60	31.4 ± 19.5	**60.0 ± 1.6**	23.2 ± 10.5	0.3 ± 0.6



(a) (b)

Fig. 2. Overlap of an US frame and its corresponding CT plane before (a) and after registration (b). The outlined tumour is from US (cyan) and CT (yellow).

The parameters in our algorithm are then estimated using the following settings: $temp_0 = 5$, $m_t = 100$, and $m = 50$.

We performed 100 registrations (20 for each target image). Table 1 shows the mean and standard deviation of the target registration error (TRE) in mm, and of the perturbed fan angle before and after registration. The error is reduced from an average of 20mm to less than 0.3mm after registration (for an isotropic voxel size of 0.5mm). This shows that the 7 parameters are estimated accurately by our method. Indeed, θ_{source} of the source CT converges to the correct value θ_{target}, regardless of its initialization, which covers a wide range of possible angles as indicated by the mean and the large standard deviation in column 2 of Table 1. These results demonstrate the ability of our method to approximate the global optimum, without being biased by the initialization.

We also show the performance of our method in the registration of real 2-D US sweeps with CT volumes of 5 patients with liver tumours. The US sweeps contained between $19 - 46$ frames and were acquired with Zonare at $10 - 12$ Hz (patients $1 - 4$), and Siemens (patient 5). The CT volumes have a slice thickness of 2.5mm. We use the following parameter settings: $temp_0 = 3$, $m_t = 200$, $m = 50$, and an initial fan angle of 20 degrees. A quantitative assessment of the registration quality is difficult in this application because of the limited output field of view. However, we report the Dice values for tumour overlap across sweeps after registration, in comparison to the ones obtained for a fan angle of 20 degrees before registration. The results in Table 2 show that our method yields a tumour overlap of more than 64%. This is a good estimate with our rigid registration method, which does not account for the non-linear deformations due to changes in tumour characteristics and position of the liver between different

Table 2. Dice before and after US-CT registrations in 5 patients

Dice (%)	Patient				
	1	2	3	4	5
before reg	12	62	53	64	74
after reg	66	68	65	64	79

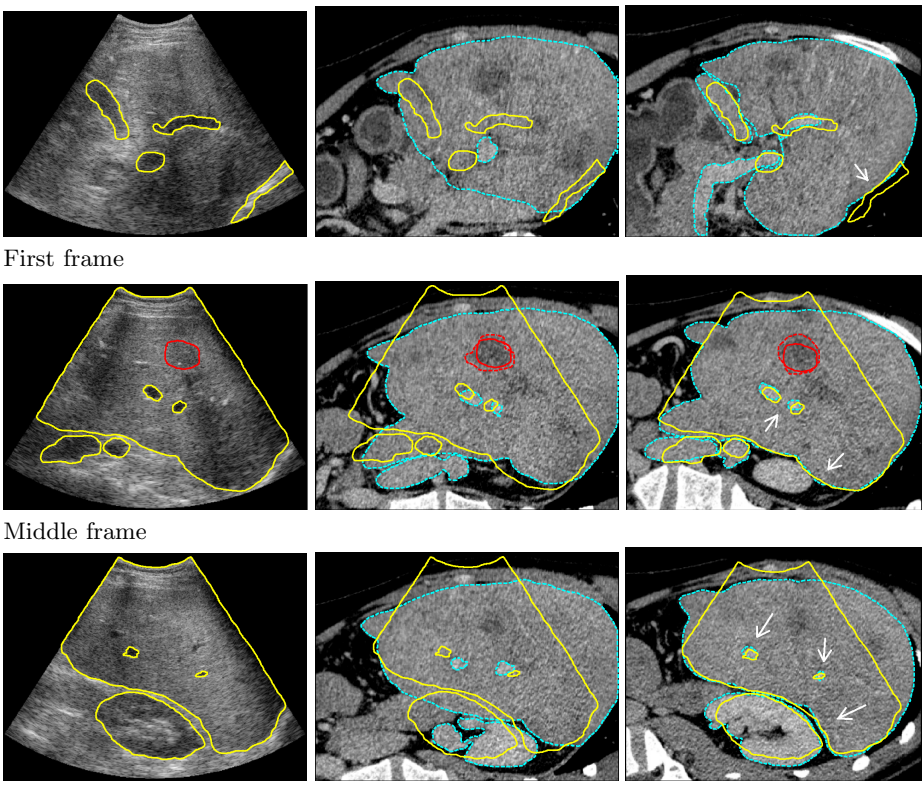


Fig. 3. First, middle and last US frames in a sweep (a) and corresponding CT planes before (b) and after (c) registration; (b) and (c) show overlaid boundaries of some visible structures in US (solid yellow contours) and CT (dashed cyan contours); the tumour is visible in both US (solid red contour) and CT (dashed red contour) in the middle frame

acquisitions. Note that the US and CT were not necessarily acquired at the same time-point, and therefore, such changes are likely to occur.

Calculating the tumour overlap before registration, e.g. using Dice, can be somewhat deceiving as it may generate high values when reslicing through the tumour at an inaccurate angle (e.g. in patients $2-5$). In this case, surrounding structures (e.g. vessels, liver boundary) will still be misaligned. For instance, Fig. 2 shows the overlap in patient 4 of an US frame and the corresponding CT plane before (a) and after (b) registration. The tumour overlap is high in both (a) and (b), which is also confirmed by the corresponding high Dice value in Table 2. However, the misalignment of the anatomical structures in (a) clearly indicates that the two scans contain different regions of the tumour (outlined in cyan in US and yellow in CT). In contrast, all structures in (b) are adequately

registered, particularly the liver boundary as pointed out by the arrow. This increases the certainty that we visualize in (b) the same region of the tumour from both scans.

We illustrate a similar case in Fig. 3, which shows three US frames in (a) from the sweep of patient 3, and the corresponding resampled CT planes before (b) and after (c) registration. In order to visually assess the ability of our method to accurately estimate θ, we show the first (row 1), middle (row 2) and last (row 3) image pairs in the sweep. Row 2 (b) and (c) exhibit high overlaps of the visible tumour outlined with solid red in US and dashed red in CT. However, the anatomical structures only correspond in (c), as pointed out by the accurate overlay of some homologous boundaries in CT (cyan) and US (yellow). Such anatomical correspondences are noticeably missing in (b). The good registration of all three frames in (c) indicates that our method estimated θ accurately.

Note that low Dice values (e.g. in patient 1 before registration in Table 2) still indicate registration failures, and the relative improvements with our method are more likely to indicate increase in registration accuracy.

In addition, we also measured a registration error from a set of landmarks identified by a clinical expert on some visible vessels in corresponding US and registered CT frames, in patients 2, 3 and 5. We found the errors to be 5.6mm in patient 2, 3mm in patient 3, and 4mm in patient 5. These values are low in this challenging untracked US-CT registration.

4 Conclusions

We presented a rigid registration framework for 2-D freehand US sweeps and 3D CT volumes, obtained without the help of probe tracking devices and phantom calibration. We defined a specific acquisition protocol, which allows to strongly constrain the registration problem. Registering N 2D US frames to a 3D CT image is indeed a particularly challenging problem, which we reduced to the estimation of only 7 parameters using simulated annealing and a robust modality-independent similarity measure. We registered the CT volume with the entire US sweep in a group-wise manner. The synthetic analysis shows that our method is robust, and the results using real datasets indicate that it has potential use in the registration of 2D US sweeps to 3D CT of liver tumours. Our future investigation consists of comparison with other state-of-the-art registration methods for tracked and untracked US-CT data. We aim at using such a technique in clinical trials for treatment monitoring of liver cancer.

Acknowledgements. This research was supported by the Wellcome Trust / EPSRC Centre of Excellence in Medical Engineering-Personalised Healthcare, WT088877/Z/09/Z. A.C., L.R. and J.A.S. also wish to acknowledge the CNRS/John Fell Oxford University Press (OUP) Research Fund. M.P.H. would like to thank EPSRC and Cancer Research UK for funding within the Oxford Cancer Imaging Centre. A.C. would like to thank Dr. Sylvia Rueda for providing the window-level code.

References

1. Penney, G.P., Joosten, K., Letteboer, M.M.J., Viergever, M.A., Niessen, W.J.: Three-dimensional reconstruction of tracked freehand ultrasound using non-rigid registration. In: IEEE ISBI, pp. 492–495 (2004)
2. Wein, W., Brunke, S., Khamene, A., Callstrom, M.R., Navab, N.: Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. Med. Image Anal. 12(5), 577–585 (2008)
3. Roche, A., Pennec, X., Malandain, G., Ayache, N.: Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information. IEEE Trans. Med. Imag. 20(10), 1038–1049 (2001)
4. Lange, T., Papenberg, N., Heldmann, S., Modersitzki, J., Fischer, B., Lamecker, H., Schlag, P.: 3D ultrasound-CT registration of the liver using combined landmark-intensity information. Int. J. CARS 4(1), 79–88 (2009)
5. Kuklisova-Murgasova, M., Cifor, A., Napolitano, R., Papageorghiou, A., Quaghebeur, G., Noble, J.A., Schnabel, J.A.: Registration of 3D fetal brain US and MRI. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part II. LNCS, vol. 7511, pp. 667–674. Springer, Heidelberg (2012)
6. Housden, R., Gee, A.H., Prager, R.W., Treece, G.M.: Rotational motion in sensorless freehand three-dimensional ultrasound. Ultrasonics 48(5), 412–422 (2008)
7. Laporte, C., Arbel, T.: Combinatorial and probabilistic fusion of noisy correlation measurements for untracked freehand 3-D ultrasound. IEEE Trans. Med. Imag. 27(7), 984–994 (2008)
8. Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, S.M., Schnabel, J.A.: Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. Med. Image Anal. 16(7), 1423–1435 (2012)
9. Cifor, A., Risser, L., Chung, D., Anderson, E., Schnabel, J.: Hybrid feature-based diffeomorphic registration for tumour tracking in 2-D liver ultrasound images. IEEE Trans. Med. Imag. PP(99) (2013)
10. Geman, S., Geman, D.: Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. 6(6), 721–741 (1984)
11. Zikic, D., Glocker, B., Kutter, O., Groher, M., Komodakis, N., Kamen, A., Paragios, N., Navab, N.: Linear intensity-based image registration by markov random fields and discrete optimization. Med. Image Anal. 14(4), 550–562 (2010)
12. Szu, H., Hartley, R.: Fast simulated annealing. Physics Letters A 122(3-4), 157–162 (1987)

Multiphase Liver Registration from Geodesic Distance Maps and Biomechanical Modelling

Jordan Bano[1,2], Stéphane A. Nicolau[1], Alexandre Hostettler[1], Christophe Doignon[2], Jacques Marescaux[1,3], and Luc Soler[1,3]

[1] IRCAD, Virtual-Surg, Place de l'Hopital 1, 67091 Strasbourg Cedex, FRANCE
[2] ICube (UMR 7357 CNRS), Boulevard S. Brant, BP 10412 67412 Illkirch Cedex, FRANCE
[3] IHU, Institut Hospitalo-Universitaire, Place de l'Hopital 1, 67091 Strasbourg Cedex, FRANCE
jordan.bano@gmail.com

Abstract. Preoperative planning for surgery is usually performed according to multiphase CT acquisitions: liver arteries and liver veins are provided from two different contrasted CT images. However, these images must be registered as they are acquired at breath hold, which are usually not identical. In this paper, we tackle this issue by providing a non-rigid registration method between the 3D liver models extracted from both preoperative images. This method is based on geodesic distance maps according to relevant landmarks and is divided in two steps: an original deformation field computation on liver surface according to geodesic distance and a biomechanical deformation of a volume mesh using our deformation field. We evaluate our method using four sets of images illustrating our clinical context. Results show that the average registration accuracy is below 1 mm for liver surface and within 5 mm for liver vessels.

Keywords: Registration, liver, multiphase, geodesic distance.

1 Introduction

Liver cancer is one of the most common cancers in the world [1]. Multiphase images are usually acquired to provide the full vessel system for preoperative surgery planning [2]. This imaging protocol consists in injecting contrast agent to obtain an image during the portal-venous phase and another one during the arterial phase. Practically, since the patient does not hold his/her breathing at the same respiratory phase during both image acquisitions, the liver and thus, its vessel, are usually deformed from one image to another. Indeed, the breathing causes an important displacement in cranio-caudal direction and deformations due to diaphragm and rib movements. Therefore, surgeons need to register mentally the complementary information provided by both images. A registration of the liver in both images would be extremely useful to help surgeons during diagnosis and preoperative planning.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 165–174, 2013.

Several authors evaluated a non-rigid registration method based on the minimization of mutual information and using a B-spline deformation field, for multiphase registration [3,4,5]. Kwon et al. [6] and Cao et al. [7] propose the use of corner features: 3D Forstner corner [6] and Harris Corner operator [7] that are automatically matched between multiphase images using intensity based criterion. They used a free-form deformation model based on B-spline [6] or thin-plate spline [7]. Since there are no common vessels between arterial and venous images, a registration based on intensity comparison only seems unsuitable due to intensity inconsistency (cf. Fig. 1).

Heldmann et al. [8] propose to use the normalized gradient fields as registration criteria with a linear elastic regularizer. Their main contribution is to provide a penalty term to avoid superimposition between veins and arteries. Drechsler et al. [9] based their work on a landmark based registration coupled with an elastic Demons registration. This step takes the distance of a voxel to the liver boundaries into account in order to avoid a false matching of non-corresponding structures inside the liver. Contrary to previous methods, these ones used a priori information to improve the deformation field inside the liver. However, their deformation is based on an elastic model which cannot properly take real biomechanical properties into account. Methods based on mechanical model such as Finite Element Model (FEM) driven by surface information can provide realistic results [10,11]. Thus, we believe that this kind of model is mandatory to estimate realistic position of liver vessels.

To our mind, the main limit of these works is that they were not properly evaluated. Finally, since all these works evaluate the multiphase registration accuracy on liver surface only, it is difficult to understand whether their approach is really relevant or not. Indeed, the registration accuracy should be assessed on the structures relevant for surgeons i.e. liver vessels and inner structures like tumours. Such an evaluation requires at least two images in the same phase which is uncommon during multiphase acquisitions.

In this paper, since the inner liver information is unreliable for intensity based registration, we propose a two step method. We firstly compute a deformation field between both liver surfaces by matching points using an original method based on geodesic distance maps (cf. Sec. 2.2). Secondly, these matchings are used with a biomechanical engine to estimate the inner liver part deformation and its vessel positions (cf. Sec. 2.3). To rigorously evaluate our method, we finally propose to use 3d+t CT images performed during venous phase, commonly used for tumour follow-up, and show the results on the liver surface and its vein network (cf. Sec. 3).

2 Method

We propose a registration of a liver source mesh M_s (with vertices P_s) to a liver target mesh M_t (with vertices P_t). Firstly, we describe how these meshes are obtained from our input data. Afterwards, we explain the computation of the matching between P_s and P_t. Finally, the biomechanical deformation of the volume mesh V_s associated to M_s is described.

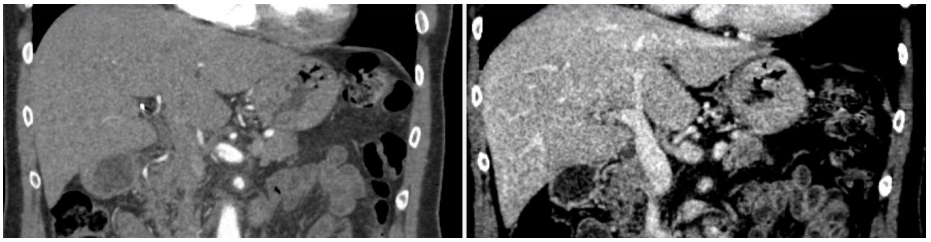


Fig. 1. Illustration of the difference of grey level intensities in multiphase liver CT-scan. Thin arteries are visible in the arterial phase (left image), whereas thick veins are visible in venous vein.

2.1 Data Input and Pre-processing

Our method needs the segmentation of the liver and its vessels on both phase images. These segmentations are performed by experts using semi-automatic tools for the preoperative planning.

Then, two surface meshes M_s and M_t (with approximately 10000 cells) are generated from the segmented images using Marching Cube algorithm. A volume mesh of the liver V_s is also computed for M_s with the CGAL library (http://www.cgal.org) (around 5000 cells) which is required for the biomechanical deformation step.

2.2 Definition of Geodesic Distance Map

The geodesic distance between two vertices on a surface is the length of the shortest path, travelling on the surface of the shape, that links the vertices. This geometric tool was used in registration of brain images [12,13], as a refinement step. However, the initial matching field was generated by another method which did not rely on geodesic distance. It also provides interesting results in the computer graphic field for isometry-invariant matching between surfaces [14]. In our case, segmentation errors occur involving that both meshes are not perfectly isometric. In this paper, we propose geodesic distance (using the geodesic library http://code.google.com/p/geodesic/) to relevant anatomical landmarks selected manually far from the segmentation errors. Our assumption is that geodesic distance between two anatomical landmarks should remain approximately constant. In fact this is not totally true since the liver can be globally compressed, which leads to a slight decrease of volume but can be easily compensated using a normalisation step.

We call a geodesic distance map GDM_I the set of geodesic distances between each mesh vertex and a seed point I. We compute a set of GDM according to selected corresponding seed points on both meshes (cf. Fig. 2). The seed points correspond to anatomically and geometrically relevant landmarks and are currently defined manually. The normalisation consists in dividing each value of a GDM_I by the maximum value of the GDM_I. To our knowledge, no work

has proposed to use the geodesic distance to landmarks as a criteria for surface matching.

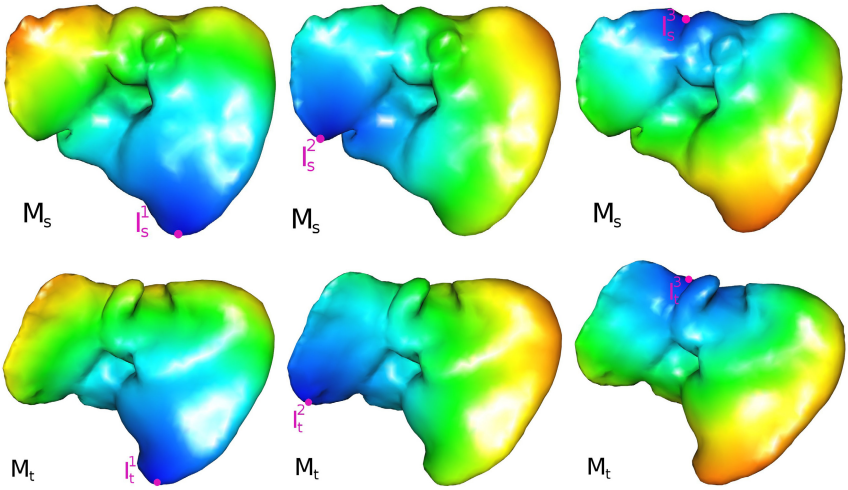


Fig. 2. One can see three different examples of geodesic distance on the liver. These meshes are coloured according to the GDM_p, with p the pink dot (blue for nearest points and red for the furthest ones). The seed points of each mesh are manually selected and matched.

2.3 Vertex Matching between M_s and M_t

The main purpose of this method is to estimate a surface deformation field from M_s to M_t which is computed by vertex matching. For each vertex P_s of M_s, we try to find its anatomical corresponding vertex on M_t.

Given two matched seed points I_s and I_t, we assume that a vertex P_t is a possible candidate match for P_s if:

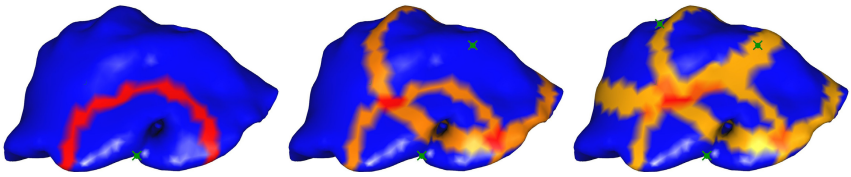
$$GDM_{I_s}(P_s) = GDM_{I_t}(P_t) \pm \varepsilon, \tag{1}$$

where $GDM_{I_s}(P_s)$ (resp. $GDM_{I_t}(P_t)$) is the geodesic distance between P_s (resp. P_t) and the seed point I_s on M_s (resp. I_t on M_t) and ε is a confidence index (which value is discussed in Sec. 3).

The main idea is that a GDM filters among all vertices of M_t the good candidates that can be matched with P_s according to Eq. 1. These vertices are distributed along a geodesic ring (which thickness depends on ε) around the seed point (cf. Fig. 3 (a)). If we consider another seed point and a supplementary GDM, we obtain another geodesic ring with another list of possible candidates. Finally, if we select the vertices which belong to both geodesic rings, the total number of candidates is then reduced (cf. Fig. 3 (b,c)).

The choice of ε is crucial and depends on the proximity between each seed point and the current vertex P_s: we assume that the information is more reliable when P_s is near the seed point than when it is far from it. Indeed, a further vertex

has a higher probability to be influenced by strong deformation or segmentation errors. Thus, the confidence index ε takes this phenomenon into account: we consider thinner ring for nearer GDM (where the information is quite sure) and larger rings for further GDM (cf. Fig. 3 (c)). We choose the ring thickness (ε in Eq. 1) equal to the geodesic distance between P_s and the GDM seed point, divided by a coefficient τ.



(a) using one GDM (b) using two GDM (c) using three GDM

Fig. 3. In this figure, we illustrate the use of multiple GDM to reduce the possible candidates P_t for one vertex P_s. The possible matchings are coloured in deep red. On the left figure, the area in red is a ring which center is the seed (green cross). On the middle figure, we add a second GDM and therefore display two areas containing candidates. On the right image, a third GDM is added and the intersection of the three rings is a small area in dark red. One can see that the thickness of the ring depends on the distance between the seed and the vertex to match P_s.

Refinement Step. When the candidate selection is finalized, we must choose one vertex P_g among the candidate set C_t only by minimizing following criteria:

$$P_g = \underset{P_k \in C_t}{\arg\min} \sum_{I_s^j, I_t^j \in SeedPoint} ||(GDM_{I_s^j}(P_s) - GDM_{I_t^j}(P_k))|| \bullet \delta,$$
$$\text{where } \delta \text{ is a normalized weight: } \delta = \frac{[GDM_{I_s^j}(P_s)]^{-1}}{\sum_{i \in [0;NbSeedPoint]} [GDM_{I_s^i}(P_s)]^{-1}} \tag{2}$$

We can also skip the filter step and do the minimization for all the vertices of M_t. But, in this case, the computation time is higher (one minute) whereas with the filter step, our method takes a few seconds only.

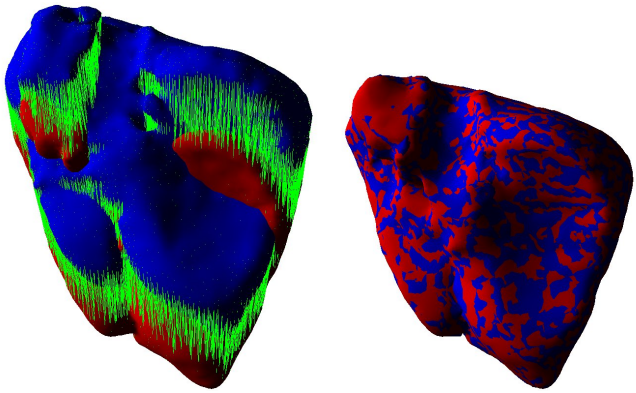
Seed Points. The seed points correspond to anatomically and geometrically relevant landmarks and are defined manually to avoid any errors of matching between them. Ideally, this step of seed selection on both meshes should be automatic but this is not the scope of this paper.

2.4 Biomechanical Deformation

The resulting matches provide a displacement field for the external surface of the liver. A biomechanical model is used to interpolate this field on the inner liver part. The volume mesh V_s of the source mesh is associated with a finite element model (a geometrically non-linear elastic law and a co-rotational formulation)

for soft tissue deformation. M_s and this volume mesh are mapped together: each vertex of M_s is associated with a tetrahedron of the V_s. Thus, if a displacement is applied on a vertex of M_s, a corresponding displacement is propagated to the associated tetrahedron of V_s using the transpose of the Jacobian of the mapping. In a same way, the vessel mesh is mapped with V_s: when V_s is deformed the vessels are also deformed. The deformation of M_s is performed by adding springs between the matched points with an empirically selected stiffness so that M_s overlaps M_t (cf. fig. 4).

The biomechanical parameters used for the finite element model are the Young's modulus and Poisson ratio. These parameters represent the deformation and compressibility properties of the liver. We choose realistic values found in literature for the Young's modulus (15 kPa found in [15]). The Poisson ratio is equal to 0.35 to allow slight volume compression or dilatation. Indeed, it happens that the volume of the liver slightly changes depending on the breath hold.



(a) Before deformation (b) Final result

Fig. 4. One can see the two meshes before and after our biomechanical deformation using SOFA. The source mesh M_s is in blue and each of its vertices are pulled to reach its matched vertex on the target red mesh M_t using springs (in green).

3 Evaluations on Human Data

We have evaluated our work on four patient cases using 3D+t venous CT. Indeed, multiphase imaging does not provide the same vessel systems so we cannot evaluate the quality of the vessel registration using this kind of acquisitions. Moreover, 3d+t images with a time gap of one year practically causes more deformations that breathing only. Thus, if our method works with these images, it should work on multiphase images. These CT have on average spatial resolution of 0.8 x 0.8 x 1 mm and an image size of 512 x 512 x 400 slices.

For the first, the second and the third patient, we had two CT with different breathing with a time gap of one year. For the third patient, we had one CT in supine position and one CT on decubitus lateral position and a time gap of two years. This data can demonstrate that even strong deformations can be

tackled by our approach. We evaluate our registration accuracy by measuring the distance between the resulting meshes of the liver surface M_r and vessel and those of the target mesh M_t.

For our experiment, we use five to ten seed points. We use a home-made software to select these vertices by picking on meshes. The user is guided by a colour scheme to pick sparse seed points. The number of seed points depends on the liver shape: it usually allows easy identification and matching of at least five anatomical points, which is already sufficient for our method to work. In case more anatomical points can be easily matched, we use all of them since it should improve the matching quality. In order to evaluate the variability of our method to the seed selection, we compute on each patient the registration using different seed pair set selected by eight different users.

The parameter τ used during the filter step is chosen empirically to 5 so that the ring thickness is larger than the minimal size of the mesh triangle. The seed point picking usually takes two to three minutes and the automatic matching step needs few seconds. Finally, the biomechanical deformation is performed by SOFA framework (http://www.sofa-framework.org/) in thirty seconds to one minute. It mainly depends on the volume mesh size (5000 cells were sufficient in our experiments). Finally, the whole registration requires less than five minutes.

3.1 Evaluation of the Mesh Surface Position

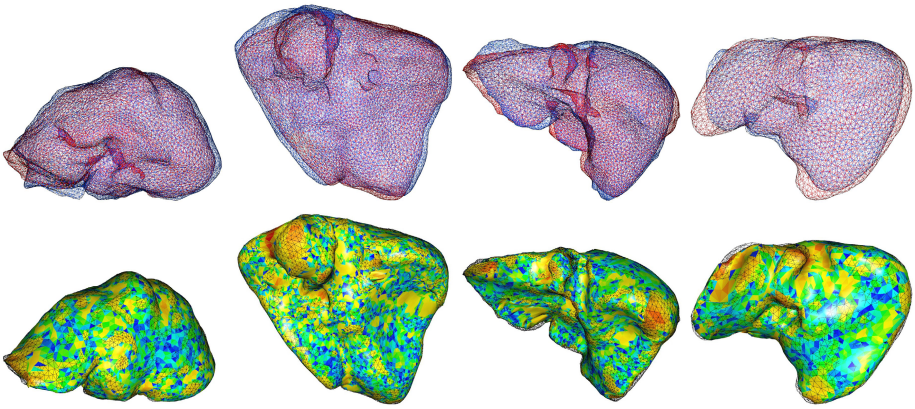
We compare our computed surface mesh M_r with the target mesh M_t. We provide a colour scheme for M_r which illustrates the distance between it and M_t. This colour scheme is done by computing the distance between each triangle T_r of M_r and the mesh M_t. The distance chosen is the length of the orthogonal projection of the gravity center G of T_r on the closest triangle of M_t. This triangle is the one whose gravity center is the closest to G. The contribution of each triangle is weighted according to its area size. We obtain a mean error within 0.5 mm in the four cases and a maximum error of 7 mm (cf. Fig. 5). The maximum errors are due to local deformations or due to segmentation mistakes. For reference, we rigidly register M_s with M_t using an ICP method and the average distance is between 2 and 4 mm (max. 10 to 20 mm). Note that the reported values are in fact an average of the results obtained with the eight different seed set selected by the eight users. The inter-user variability has been computed and is below 10%, which means our method is sufficiently user independent.

3.2 Evaluation of the Vessel System

The evaluation of the vessel registration accuracy is performed by computing the euclidean distance between ten vein bifurcations which have been manually selected (cf. Tab. 2 and Fig. 6). We obtain an average error of 2.8 mm instead of 4.6 mm, which is sufficient according to our surgeons. Although the rigid registration seems to provide initial good results, we highlight that the main error is usually due to a lobe which has been highly deformed (5 mm to 15 mm). Thus, our method provides the strongest improvement (corresponding to

Table 1. The distance between each triangle of M_R and the ground truth segmentation is computed and sorted in four groups. Each group contains a quartile of the triangle total number weighted by its area size. The distance is between M_R and M_T: after rigid ICP registration (1) and after non-rigid registration (2).

	First case (mm)		Second case (mm)	
Colour range	(1)	(2)	(1)	(2)
Blue to Turquoise	0 - 0.8	0 - 0.1	0 - 1.0	0 - 0.05
Turquoise to Green	0.8 - 1.8	0.1 - 0.2	1.0 - 1.8	0.05 - 0.1
Green to Yellow	1.8 - 3.5	0.2 - 0.5	1.8 - 2.9	0.1 - 0.2
Yellow to Red	3.5 - 13.3	0.5 - 6.3	2.9 - 10.3	0.2 - 4.7
Mean (± Std.Dev.)	1.9 (± 1.9)	0.3 (± 0.4)	1.4 (± 1.7)	0.1 (± 0.3)
	Third case (mm)		Fourth case (mm)	
Colour range	(1)	(2)	(1)	(2)
Blue to Turquoise	0 - 1.7	0 - 0.1	0 - 1.8	0 - 0.1
Turquoise to Green	1.7 - 4.1	0.1 - 0.2	1.8 - 4.0	0.1 - 0.3
Green to Yellow	4.1 - 6.1	0.2 - 0.5	4.0 - 6.6	0.3 - 0.6
Yellow to Red	6.1 - 13.2	0.5 - 6.7	6.6 - 20.6	0.6 - 5.7
Mean (± Std.Dev.)	3.2 (± 3.0)	0.3 (± 0.6)	4.0 (± 3.3)	0.5 (± 0.6)



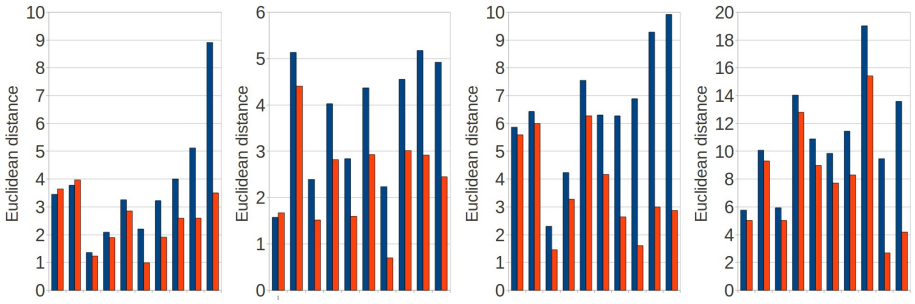
(a) First patient (b) Second patient (c) Third patient (d) Fourth patient

Fig. 5. Top: four examples of human liver pair segmented from 3d+t images. An ICP method was used to rigidly register the liver pair to underline the deformations between the meshes only. Bottom: results of our surface non-rigid registration coloured according to Table 1 with the ground truth (M_t) in black wireframe.

right columns in all histograms in Fig. 6) for lobe motion that has been properly registered (3 mm). For the fourth patient the average distance is more important than others (8 mm). We recall that for this patient we have a CT in decubitus lateral position and the other one in supine position which involves strongest deformations. Despite an insufficient accuracy, our method have a promising average improvement of 160% for this patient.

Table 2. Vessel registration accuracy of both patients according to the euclidean distance of ten vessel bifurcations (more details in Fig. 6)

	Before registration (mm)		After registration (mm)		Average
	Average error	Max. error	Average error	Max. error	improvement (mm)
First patient	3.6 ± 2.5	8.9	2.2 ± 1.1	4.0	1.2
Second patient	3.7 ± 1.3	5.2	2.4 ± 1.1	4.4	1.3
Third patient	6.5 ± 2.2	9.9	3.7 ± 1.8	6.3	2.8
Fourth patient	11.2 ± 3.7	19.1	7.9 ± 3.8	15.4	3.0



(a) First patient (b) Second patient (c) Third patient (c) Fourth patient

Fig. 6. One can see the euclidean distance between vessel bifurcations after rigid registration (in blue) and after non-rigid registration (in orange). Each column represents the error of one bifurcation. Strongest improvement (corresponding to right columns in all histograms) corresponds to a lobe motion that has been compensated by our method.

4 Conclusion

In this paper, we have presented a novel method to register two meshes according to geodesic distance maps and using a biomechanical engine. Firstly, a vertex matching is computed by filtering vertices and minimizing a selection criteria based on the geodesic distance to selected anatomical landmarks. Then, a biomechanical deformation is applied on the liver and its vessels. The evaluation of our registration gives promising results: 0.5 mm for liver surface position and 3.5 mm for vessel position. Moreover, we showed how the use of 3d+t CT for patient tumour follow-up allows to rigorously evaluate multiphase registration methods on inner liver landmarks.

We highlight our method is not only relevant to register internal structures of two livers in multiphase CT, but more generally two livers in different modality where inner grey level intensities do not provide the same anatomical information. Our method have given promising results for intra-operative liver registration in laparoscopic surgery [16]. In this context, we have shown that the liver deformation due to breathing and to pneumoperitoneum was reduced from 6 mm to 4 mm for the whole liver surface. In the future, we will develop an automatic method for seed point detection and matching using features as those proposed in [17].

References

1. Garcia, M., et al.: Global cancer facts & figures 2007, vol. 1(3). American Cancer Society, Atlanta (2007)
2. Soler, L., et al.: Computer-assisted operative procedure: from preoperative planning to simulation. European Clinics in Obstetrics and Gynaecology 2(4), 201–208 (2006)
3. Lange, T., et al.: Registration of different phases of contrast-enhanced ct/mri data for computer-assisted liver surgery planning: Evaluation of state-of-the-art methods. The International Journal of Medical Robotics and Computer Assisted Surgery 1(3), 6–20 (2005)
4. Zhu, Z., et al.: Non-rigid registration of multiphase contrast-enhanced abdomen ct image. In: 2011 6th International Conference on Computer Sciences and Convergence Information Technology (ICCIT), pp. 909–914. IEEE (2011)
5. Zhu, W., Nicolau, S., Soler, L., Hostettler, A., Marescaux, J., Rémond, Y.: Fast segmentation of abdominal wall: Application to sliding effect removal for non-rigid registration. In: Yoshida, H., Hawkes, D., Vannier, M.W. (eds.) Abdominal Imaging 2012. LNCS, vol. 7601, pp. 198–207. Springer, Heidelberg (2012)
6. Kwon, D.: et al.: Efficient feature-based nonrigid registration of multiphase liver ct volumes. In: BMVC 2008 (2008)
7. Cao, Y., et al.: An efficient feature-based non-rigid registration of multiphase liver ct images using matching region partition. In: 2010 The 5th IEEE Conference on Industrial Electronics and Applications, pp. 1657–1661. IEEE (2010)
8. Heldmann, S., et al.: Elastic registration of multiphase ct images of liver. In: SPIE Medical Imaging, p. 72591H. International Society for Optics and Photonics (2009)
9. Drechsler, K., et al.: Multiphase risk assessment of atypical liver resections. In: 2012 25th International Symposium on Computer-Based Medical Systems (CBMS), pp. 1–4. IEEE (2012)
10. Samani, A., et al.: Biomechanical 3-d finite element modeling of the human breast using mri data. IEEE Transactions on Medical Imaging 20(4), 271–279 (2001)
11. Cash, D.M., et al.: Compensating for intraoperative soft-tissue deformations using incomplete surface data and finite elements. IEEE Transactions on Medical Imaging 24(11), 1479–1491 (2005)
12. Wang, Y., et al.: 3d brain surface matching based on geodesics and local geometry. Computer Vision and Image Understanding 89(2), 252–271 (2003)
13. Cho, Y., et al.: A multi-resolution scheme for distortion-minimizing mapping between human subcortical structures based on geodesic construction on riemannian manifolds. NeuroImage 57(4), 1376–1392 (2011)
14. Bronstein, A., et al.: Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. Proceedings of the National Academy of Sciences of the United States of America 103(5), 1168–1172 (2006)
15. Samur, E., et al.: A robotic indenter for minimally invasive characterization of soft tissues. In: International Congress Series, vol. 1281, pp. 713–718 (2005)
16. Bano, J., et al.: Registration of preoperative liver model for laparoscopic surgery from intraoperative 3d acquisition, pp. 201–211 (to appear, 2013)
17. Boyer, E., Bronstein, A.M., Bronstein, M.M., Bustos, B., Darom, T., Horaud, R., Hotz, I., Keller, Y., Keustermans, J., Kovnatsky, A., et al.: Shrec 2011: robust feature detection and description benchmark. In: Proceedings of the 4th Eurographics Conference on 3D Object Retrieval, pp. 71–78. Eurographics Association (2011)

Fast Renal Cortex Localization by Combining Generalized Hough Transform and Active Appearance Models

Dehui Xiang[1,2], Xinjian Chen[2], and Chao Jin[2]

[1] School of Computer Science and Technology, Soochow University, No.1 Shizi Street, Jiangsu 215006, China
[2] School of Electronics and Information Engineering, Soochow University, No.1 Shizi Street, Jiangsu 215006, China
xiangdehui@suda.edu.cn

Abstract. Automatic localization of objects is one of great important steps in object recognition and analysis, such as segmentation, registration in many medical applications. In this paper, an automated method is proposed to recognize renal cortex on contrast-enhanced abdominal CT images. The proposed method is based on a strategic combination of the Generalized Hough Transform and Active Appearance Model. It consists of two main phases: training and localization. In the training phase, we train the mean shape models of renal cortex by using Active Appearance Model and compute Generalized Hough Transform parameters. In the localization phase, a modified Generalized Hough Transform algorithm is advanced to estimate potential center of gravity for improving the conventional Active Appearance Model matching method, and then a two-pass Active Appearance Model matching method is proposed based on Generalized Hough Transform. The Active Appearance Models and Generalized Hough Transform parameters were trained with 20 CT angiography datasets, and then the proposed method was tested on a clinical data set of 17 CT angiography datasets. The experimental results show that: (1) an overall cortex localization accuracy is 0.9920 ± 0.0038, average distance is 11.00 ± 9.34 pixels. (2) The proposed method is highly efficient such that the overall localization can be finalized within 1.2075 ± 0.3738 seconds for each 2D slice.

Keywords: Localization, kidney, renal cortex, generalized hough transform, active appearance model.

1 Introduction

Kidney cancer is among the 10 most common cancers worldwide in both men and women. It is investigated that the risk for developing kidney cancer is about 1/70 in the lifetime [1]. Renal cell carcinoma is the most common type of kidney cancer in adults (approximately 80% of cases), on the other hand, it is studied that renal cell carcinoma mainly arises from the renal cortex [2]. Therefore, the

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 175–183, 2013.

research of the renal cortex has great value for kidney cancer diagnosis and treatment.

Recognition of kidney and renal cortex is of particular clinical importance. It can be helpful to accurately delineate its specific anatomical structures, such that physicians and radiologists can diagnose various disorders, locate pathologies, create statistical atlases, quantify structural properties, etc. However, kidney and renal cortex recognition in many medical situations is still based mainly on manual methods by radiologists. Those traditional methods are usually subjective, tedious, and prone to errors in clinical applications. Therefore, a fully automatic and accurate method is needed to accurately recognize and analyze kidney and renal cortex. The current methods may be classified into two types: unsupervised methods and supervised methods for recognition and analysis. Unsupervised recognition methods, which are often intensity-based (e.g., thresholding, region growing [3], morphological operations [4], clustering), usually not too accurate in recognition and analysis of renal cortex, since there is much biological variability in appearance of both normal and pathological anatomical structures, image noise and artifacts from imaging devices. On the other hand, supervised recognition methods, which are usually model-based (e.g., atlases [5], statistical active shape models [6], and statistical active appearance models [7–9]), can better deal with such variable conditions by integrating prior knowledge of anatomical structures into the recognition and analysis procedure. In the case of landmark-based supervised recognition and analysis, the intensity appearance and spatial relationships of region of interest can be modeled by locating landmarks at corresponding anatomical positions of renal cortex in a series of representative CT or MRI images during the process of learning, such that prior knowledge can be obtained and then used to recognize and analyze renal cortex in target images.

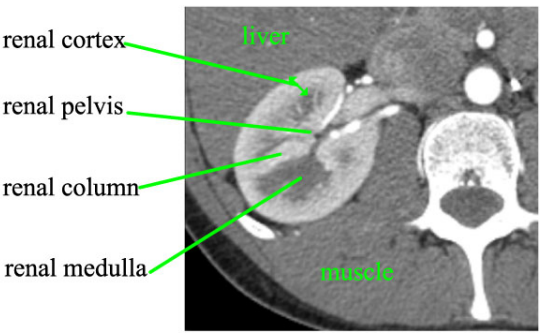


Fig. 1. The illustration for anatomy of kidney

In order to accurately recognize and analyze renal cortex by supervised recognition methods, an effective and efficient initialization method is needed to locate kidney or renal cortex shape in the test images. Initial localization of renal cortex is not a trivial task. The anatomical structures of kidneys are complex since

there are four major internal structures, the renal cortex, renal column, renal medulla, and renal pelvis, as shown in Fig.1. The renal cortex and renal column are connected and have the similar intensity distributions. In addition, there are usually blurred boundaries between kidneys and adjacent organs such as liver, blurred boundaries between renal cortex and renal medulla. There were several prior investigations in renal cortex localization for renal cortex recognition and analysis in CT and MRI images, including both semi-automatic and fully automatic methods. A semi-automatic method was presented by Shen et al. [10] based on the morphological 3D h-maxima transform to locate the kidneys in MR images. Meanwhile, a temporal Markov model was introduced by Boykov et al. [11] to compute the time intensity curves, and then the graph cut algorithm was used to segment kidneys, however, foreground seeds and back-ground seeds need to be painted by users for the localization. Although the performance was in clinically acceptable level, specifying the seed points for the segmentation made it practically not useful due to sensitivity of seed localizations. For the atlas based and registration based methods, an image registration algorithm was proposed by Sun et al. [12] to locate renal cortex for dynamic renal perfusion MR images. Another registration based approach was proposed by Zollner et al. [13] to separate the inner components based on k-means.

In this paper, we propose a framework to automatically locate renal cortex precisely and quickly. In the rest of the paper, in Section 2, we elaborate the complete methodology of the localization algorithm. In Section 3, we describe an evaluation of this method in terms of its accuracy and efficiency. In Section 4, we summarize our contributions and conclusions.

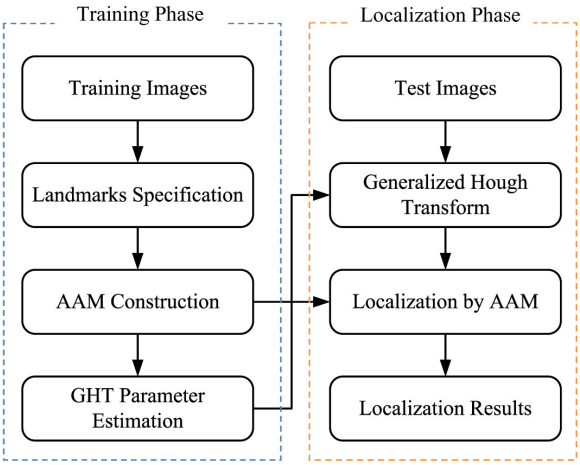


Fig. 2. The flowchart of the proposed method

2 Method

The proposed framework consists of two phases: training phase and localization phase. Fig.2 shows the flowchart of the proposed method. In the training phase, we construct the Active Appearance Model and compute Generalized Hough Transform parameters. This allows us to construct a mean renal cortex shape and a reference table of the mean kidney shape. The localization phase consists of two main parts: Generalized Hough Transform and localization. For the Generalized Hough Transform part, a modified Generalized Hough Transform algorithm is advanced to estimate potential center of gravity (CoG) for improving the conventional Active Appearance Model matching method. For the localization part, an Active Appearance Model matching method is proposed based on the proposed Generalized Hough Transform method.

2.1 Model Building and Parameter Training

In this phase, renal cortex were manually labeled slice by slice in axial view, although semi-automatic or automatic methods are also available [14]. The model building requires several constraints, one of which is to define anatomical correspondence. For each axial slice, the candidate landmarks were visualized dynamically, and then were identified in a clockwise direction on the outer and inner boundaries of renal cortex respectively. Since the size and the location of renal cortex may vary considerably from one patient to another, it would be desirable to interpolate the image slices of renal cortex according to remarkable physical location correspondences between patients. Therefore, we first manually distinguished the top and bottom slices of renal cortex. Then the same number of slices was obtained by linear interpolation in the training dataset. Subsequently, 2D mean shape models based on Active Appearance Model are then constructed for each axial slice level from the training images. In order to compute the directions of mean shape models, The edges and directions of mean renal cortex models are detected through Sobel edge detector. Since the regions of inner

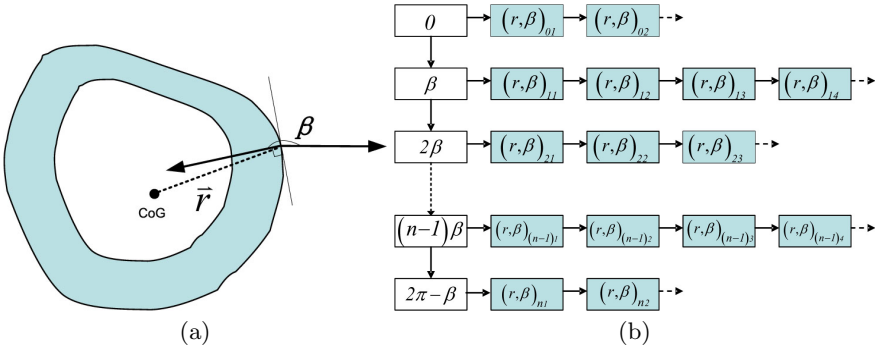


Fig. 3. (a) Generalized Hough Transform; (b) R table

contours may contain several compartments and the boundaries between renal parenchyma and renal medulla, as shown in Fig.1, the outer edges of mean renal cortex models was used to compute Generalized Hough Transform parameters. The outer edges of mean renal cortex models are stored in a reference table (R-table) format in the Generalized Hough Transform [15]. The center of gravity of outer edges is computed to consider as the corresponding reference point shown in Fig.3(a), and the length and direction of a vector connecting the boundary pixel to the reference point are computed for every pixel on the renal cortex boundary. The gradient directions, β, serve as indices in the R-table to look up the length $\vec{r} = \{r_x, r_y\}$ of the line, which connects the boundary pixel to the reference point. Fig.3(b) illustrates a general form of a 2D reference table.

2.2 Renal Cortex Localization

The proposed renal cortex localization method was based on the Generalized Hough Transform and Active Appearance Model. The conventional Active Appearance Model matching method was based on the root mean square intensity difference between the target/test image slice and mean appearance model, which involved to change the affine transformation, global intensity parameters, and appearance coefficients in a long range [16]. And then a gradient descent method was applied to minimize the difference between the target image and synthesized model. It takes a large amount of running time to obtain better accuracy of matching. To overcome these obstacles, we integrate the modified Generalized Hough Transform for fast renal cortex localization.

After the 2D reference table was constructed, the outer edge of mean renal cortex models are detected through Sobel edge detector while pruning edges that definitely do not belong to the renal cortex boundary. As the renal cortex has a CT value and gradient magnitude value range, the edges with corresponding CT value and gradient magnitude value out of the range will be removed for an accurate and fast Generalized Hough Transform.

When detecting renal cortex in a target image, the gradient angle β of an edge point is employed to retrieve corresponding entries of the R-table. An accumulator image for parameter space saves the votes of edge points to determine the most probable CoG of renal cortex outer layer. In order to reduce computation complexity in a target image, we only restrict the mean shape model transform under translation and isotropic scaling. The experiments show that the operation can reduce computational cost while achieve an acceptable localization result. For an edge point $\vec{p} = \{p_x, p_y\}$ in a target image, the corresponding $\vec{r}$ can be determined by gradient direction of the edge point. Then, the possible location of reference point $\vec{c} = \{c_x, c_y\}$ in the parameter space is calculated from $\vec{c} = \vec{p} + s\vec{r}$, where s denotes the scaling factor. When the possible reference location is computed, the corresponding pixel value in accumulator image $O\left(\vec{c}\right)$ needs +1 increments. In order to reduce artifacts and spurious spots, the regularized Generalized Hough Transform can be computed by

$$O_n\left(\vec{c}\right) = \begin{cases} O\left(\vec{c}\right), O\left(\vec{c}\right) \geq T; \\ 0, \qquad else. \end{cases} \tag{1}$$

$$H\left(\vec{c}\right) = k \cdot \left(\frac{O_n\left(\vec{c}\right)}{k_n}\right)^{\alpha} \tag{2}$$

where, T is a constant threshold, k and k_n are regularized parameters, and α is a rigid parameter to enhance the candidate reference points. The Active Appearance Model searching method is performed based on the mean appearance model, the modified Generalized Hough Transform and the pose parameters (i.e., rotation and scale).

3 Experiments and Results

The proposed methods were tested on a clinical CT data set, containing contrast-enhanced 37 volume data. The data were acquired from two different types of CT scanner (GE Medical systems, Light-Speed Ultra, and Philips, Mx8000 IDT 16). The pixel varied from 0.55 to 1 mm, and slice thickness from 1 to 5 mm.

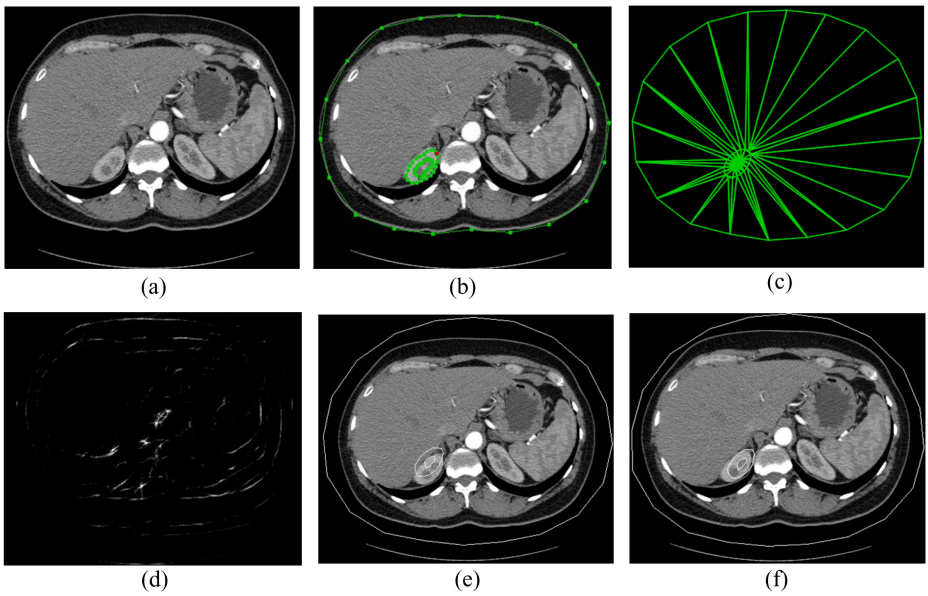


Fig. 4. Experimental result for renal cortex localization. (a) The original image; (b) The Landmark for renal cortex; (c) The corresponding mean Active Appearance Model; (d) The candidate reference points of renal cortex by using modified Generalized Hough Transform; (e) Renal cortex localization by the traditional method; (f) Renal cortex localization by the proposed method.

20 volume data were used as training data, while the other 17 volume data were considered as testing data. Renal cortex were manually segmented to generate the reference images (ground truth). Our method was implemented and tested on a 32-bit desktop PC (3.1 GHz CPU and 4 GB RAM). Fig.4 shows the localization results for renal cortex by conventional method [16] and the proposed method, respectively. For quantitative evaluation, true positive volume fraction (TPVF), false positive volume fraction (FPVF), true negative volume fraction (TNVF), false negative volume fraction (FNVF), accuracy [17], CoG distance between reference images and localization images, running time were used to evaluate the proposed method. The evaluation results are summarized in Table 1. The running time by the proposed localization method was almost as 1/2 as that of conventional method. The average TPVF, FPVF, TNVF, FNVF, accuracy and distance for renal cortex localization achieved a slight improvement.

Table 1. Quantitative comparisons between conventional method and the proposed method (Mean ± SD)

	TPVF	FPVF	TNVF	FNVF
Conventional method	0.4703±0.2018	0.0040±0.0020	0.9959±0.0020	0.5296±0.2018
Proposed method	0.4719±0.1981	0.0040±0.0019	0.9959±0.0019	0.5280±0.1981
		Distance(pixels)	Accuracy	Time(s)
Conventional method		11.2899±9.4105	0.9919±0.0038	2.3018±0.3164
Proposed method		11.0009±9.3490	0.9920±0.0037	1.2075±0.3738

4 Conclusions

In this paper, we proposed a fast automated framework for renal cortex localization. The proposed method is based on a strategic combination of the Generalized Hough Transform and Active Appearance Model. It consists of two main phases: training and localization. In the model training phase, we train the mean shape models of renal cortex by using Active Appearance Model and compute Generalized Hough Transform parameters. In the localization phase, a modified Generalized Hough Transform algorithm is advanced to estimate potential center of gravity for improving the conventional Active Appearance Model matching method. The proposed method was tested on a clinical data set of 37 CT angiography volume data. As shown in the experiments, the proposed automatic algorithm greatly reduced the time needed for the initialization. The mean localization time is about 1.2 seconds for renal cortex recognition and analysis in CTA images, which is much less than conventional Active Appearance Model localization method, and much more practically in clinical routine, while the

improvement of TPVF, FPVF, TNVF, FNVF, accuracy, CoG distance between reference images and localization images could be also achieved. Therefore, the combination of Generalized Hough Transform and Active Appearance Model will be extended to 3D formulation more precisely and efficiently.

Acknowledgments. This paper is supported by the National Basic Research Program of China (973 Program) under Grant 2014CB748600.

References

1. American Cancer Society: What are the key statistics about kidney cancer?, http://www.cancer.org/Cancer/KidneyCancer/DetailedGuide/kidney-cancer-adult-key-statistics
2. Godley, P.A., Taylor, M.: Renal cell carcinoma. Current Opinion in Oncology 13(3), 199–203 (2001)
3. Rusko, L., Bekes, G., Nemeth, G., Fidrich, M.: Fully automatic liver segmentation for contrast-enhanced ct images. In: MICCAI Wshp. 3D Segmentation in the Clinic: A Grand Challenge, vol. 2(7) (2007)
4. Fujimoto, H., Gu, L., Kaneko, T.: Recognition of abdominal organs using 3d mathematical morphology. Systems and Computers in Japan 33(8), 75–83 (2002)
5. Park, H., Bland, P.H., Meyer, C.R.: Construction of an abdominal probabilistic atlas and its application in segmentation. IEEE Transactions on Medical Imaging 22(4), 483–492 (2003)
6. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision 1(4), 321–331 (1988)
7. Frangi, A.F., Rueckert, D., Schnabel, J.A., Niessen, W.J.: Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling. IEEE Transactions on Medical Imaging 21(9), 1151–1166 (2002)
8. Chen, X., Summers, R.M., Cho, M., Bagci, U., Yao, J.: An automatic method for renal cortex segmentation on ct images: Evaluation on kidney donors. Academic Radiology 19(5), 562–570 (2012)
9. Chen, X., Udupa, J.K., Bagci, U., Zhuge, Y., Yao, J.: Medical image segmentation by combining graph cuts and oriented active appearance models. IEEE Transactions on Image Processing 21(4), 2035–2046 (2012)
10. Shen, W., Kassim, A.A., Koh, H., Shuter, B.: Segmentation of kidney cortex in mri studies: a constrained morphological 3d h-maxima transform approach. International Journal of Medical Engineering and Informatics 1(3), 330–341 (2009)
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. Sun, Y., Jolly, M.P., Moura, J.: Integrated registration of dynamic renal perfusion mr images. In: 2004 International Conference on Image Processing, ICIP 2004, vol. 3, pp. 1923–1926. IEEE (2004)
13. 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. Computerized Medical Imaging and Graphics 33(3), 171–181 (2009)

14. Davies, R.H., Twining, C.J., Cootes, T.F., Waterton, J.C., Taylor, C.J.: A minimum description length approach to statistical shape modeling. IEEE Transactions on Medical Imaging 21(5), 525–537 (2002)
15. Khoshelham, K.: Extending generalized hough transform to detect 3d objects in laser range data. In: Proceedings of ISPRS Workshop on Laser Scanning, LS 2007, pp. 206–210 (2007)
16. Stegmann, M.B., Ersboll, B.K., Larsen, R.: Fame-a flexible appearance modeling environment. IEEE Transactions on Medical Imaging 22(10), 1319–1331 (2003)
17. Fawcett, T.: An introduction to roc analysis. Pattern Recognition Letters 27(8), 861–874 (2006)

3D Surface Reconstruction of Organs Using Patient-Specific Shape Priors in Robot-Assisted Laparoscopic Surgery

Alborz Amir-Khalili[1], Jean-Marc Peyrat[2], Ghassan Hamarneh[3], and Rafeef Abugharbieh[1]

[1] Biomedical Signal and Image Computing Lab, University of British Columbia, 2329 West Mall, Vancouver, Canada
alborza@ece.ubc.ca
[2] Qatar Robotic Surgery Centre, Qatar Science and Technology Park, Education City, Doha, Qatar
[3] Medical Image Analysis Lab, Simon Fraser University, 8888 University Drive, Burnaby, Canada

Abstract. With the advent of robot-assisted laparoscopic surgery (RALS), intra-operative stereo endoscopy is becoming a ubiquitous imaging modality in abdominal interventions. This high resolution intra-operative imaging modality enables the reconstruction of 3D soft-tissue surface geometry with the help of computer vision techniques. This reconstructed surface is a prerequisite for many clinical applications such as image-guidance with cross-modality registration, telestration, expansion of the surgical scene by stitching/mosaicing, and collision detection. Reconstructing the surface geometry from camera information alone remains a very challenging problem in RALS mainly due to a small baseline between the optical centres of the cameras, presence of blood and smoke, specular highlights, occlusion, and smooth/textureless regions. In this paper, we propose a method for increasing the overall surface reconstruction accuracy by incorporating patient specific shape priors extracted from pre-operative images. Our method is validated on an *in silico* phantom and we show that the combination of both pre-operative and intra-operative data significantly improves surface reconstruction as compared to the ground truth. Finally, we verify the clinical potential of the proposed method in the context of abdominal surgery in a phantom study of an *ex vivo* lamb kidney.

Keywords: Surface reconstruction, computational stereo, shape prior, robot-assisted minimally invasive surgery.

1 Introduction

The term minimally invasive laparoscopic surgery refers to interventional techniques (performed in abdominal and pelvic cavities through small key-hole incisions) that reduce trauma and morbidity as opposed to traditional open surgery. In these procedures, surgeons access the surgical site using an endoscopic camera

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 184–193, 2013.

and miniaturized (laparoscopic) instruments such as graspers, scissors, suction, and electro-cautery tools. Traditional laparoscopic instruments suffer from difficult hand-eye ergonomics and a restricted 2D field of view of the surgical site acquired via the endoscopic camera. New RALS technologies, da Vinci Surgical System (Intuitive Surgical, Inc., Sunnyvale, CA, USA), have overcome many of these restrictions by incorporating instruments with six degrees of freedom, thus improving tissue manipulation dexterity, and stereo endoscopic cameras that afford better depth perception of surgical site.

These robots facilitate the next generation of image guided interventions in abdominal surgery by establishing a natural interface between the surgeon, robot, and patient. Advanced techniques in computer vision can now be integrated alongside the stereo endoscopic cameras of such systems to enhance the visualization capabilities of the surgeon. Three dimensional (3D) reconstruction [1] of soft-tissue surfaces from stereo cameras is an essential intermediary step used within many image guidance techniques. As surveyed in a recent publication [2], such laparoscopic applications include: augmented reality, telestration, feature tracking, dynamic view expansion, biophotonics, and collision detection.

According to [2], many hardware solutions, e.g. structured light and time of flight, and software algorithms, e.g. shape-from-motion and SLAM, have been proposed for 3D surface reconstruction of abdominal organs. However, computational stereo (reconstruction from dense matching) may be the strongest candidate for translation into clinical practice since: 1) stereo-endoscopic imaging hardware is readily available with the da Vinci system, and 2) it produces a more accurate and dense reconstruction compared to other software solutions. Computational stereo methods are, however, prone to error in presence of non-Lambertian specular highlights, smoke, blood, and tissue surfaces that are smooth (Fig. 1). These problem areas lack the definitive features that are essential for dense matching. Furthermore, the accuracy of depth reconstruction drops dramatically if the distance between the target tissue and a camera center is more than 15 times the baseline which is the physical distance between the optical centers of the cameras. Due to the small size of the endoscope, the baseline is typically very small (around 5 mm).

The afforementioned errors are especially prevalent during surface reconstruction of the kidneys and liver, both of which are round, i.e. have high extrinsic curvature, *and* are smooth in terms of shape and texture. In this paper we focus mainly on the kidney since delicate procedures such as radical/partial nephrectomy and pyeloplasty are now being performed with da Vinci systems [3].

To overcome such errors and limitations during RALS, we are proposing to incorporate geometric shape priors derived from patient-specific computed tomography (CT) scans acquired prior to surgery as a regularization constraint on top of a robust dense matching algorithm. Provided that the prior is accurately registered to the endoscopic view, the number of outliers (mismatched features on the surface of the organ) can be drastically reduced without oversmoothing the discontinuties that exist.

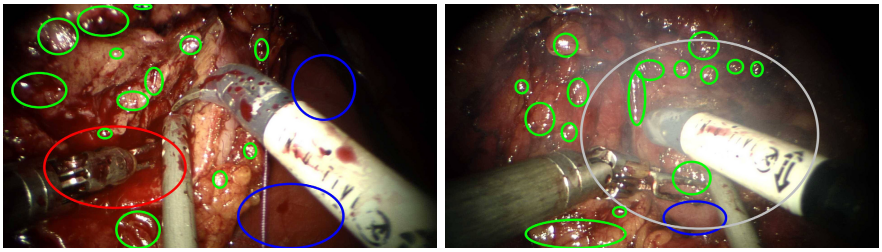


Fig. 1. Example of "artefacts" present in images acquired by the left stereo-endoscopic camera of the da Vinci robot during a partial nephrectomy. These "artefacts" can induce error in the stereo reconstruction methods based on dense feature matching: blood (red), white specular highlights (green), smooth organ surfaces (blue), and smoke (gray).

After describing the details of our method in Section 2, we perform validation on an *in silico* cardiac phantom [4,5]. To the best of our knowledge, this is the only publicly available dataset captured with the da Vinci System that includes ground truth surface reconstruction. In this study, we demonstrate a significant improvement by comparing a basic reconstruction algorithm, a more robust algorithm, and our proposed reconstruction with shape prior encoding method to the available ground truth. An *in silico* cardiac dataset is not sufficient to justify the use of our method for its intended application in abdominal surgery. To suppliment our proof of concept, we apply our method to an *ex vivo* lamb kidney, captured with the da Vinci Surgical System, to qualitatively demonstrate its potential utility in a clinical setting.

2 Methods

In this section we present two approaches of reconstructing depth information from stereo endoscopic cameras followed by our proposed method which includes a shape prior. First, a basic method is presented and used as a reference to measure improvements in reconstruction accuracy. The second method, is more robust and similar to state-of-the-art methods that do not use shape priors. Our proposed method is an extension of the robust method with the addition of manually posed patient-specific shape prior extracted from CT.

2.1 Basic Stereo-Endoscopic Surface Reconstruction

We assume that adequate calibration of the stereo rig is performed and the camera projection matrices of the pinhole camera model are available. We take advantage of 3D reconstruction methods from multi-view geometry [1] to simplify the problem and avoid a direct 3D dense matching. We simplify the dense

matching problem by transforming the pair of views using a polar rectification method [6]. This reduces the dense matching to linear matching along parallel epipolar lines in the left and right rectified images from which we obtain correspondences.

First, the rectified images are converted into greyscale images. We use the normalized cross correlation (NCC) ratio as the similarity metric. This matching criteria has the advantage of being less prone to changes in illumination. NCC is computed for matching patches of pixels centered at the points being matched in each image. A square patch is used for matching and set to 25×25 pixels, a patch size emperically shown to provide a good compromize between the number of outliers and oversmoothness of the resulting depth values.

In this basic approach, the matching points with the highest NCC value are selected. The corresponding distance between matches (in pixels) in the unrectified images is converted to position coordinates (in millimeters) in 3-space with the origin at the left camera centre. This is achieved using the calibrated camera projection matrices and a simple triangulation method using singular value decomposition [1].

2.2 Robust Surface Reconstruction

Certain constraints may be imposed to improve the accuracy of the reconstructed positions in 3-space by removing outliers that occur during matching. As previously proposed in relevant publications [7,8], the ideal matching of a patch from the left camera to the right camera must be consistent with matching the same patch from the right camera to the left. This doubles the computational complexity of the algorithm but introduces robustness to specular noise and occlusions.

The NCC metric does not guarantee a perfect match, especially if the textural information within that patch is not unique. Rather than increasing the size of the patch, the information of the neighbouring matches can be used to remove false matches. Discontinuities thus may exist in the reconstructed depth, e.g. sharp edges and occlusion of view by laparoscopic tools, hence we account for isolated mismatches with a median filter as an edge-preserving outlier removal method.

2.3 Surface Reconstruction with a Shape-Prior

We propose to further improve the robust surface reconstruction, denoted I^{RR}, by adding more constraints based on the geometry extracted from a prior model. This model may be obtained by segmenting a patient specific pre-operative image. In our examples we use manually segmented CT volumes as they are commonly acquired prior to most abdominal RALS.

The extracted shape prior is registered to the endoscopic view by manually finding corresponding points between the left and right camera images, and the prior. The calibrated camera projection matrices are used to automatically pose the CT onto both cameras using a similarity transform consisting of rigid

motion and uniform scaling. By posing the shape prior onto the CT, we are able to project its surface onto the camera views using the camera matrices. This projection results in a depth reconstruction, which is in the same domain as I^{RR}, hereby denote I^{CT}.

The shape prior of the CT is encoded into the reconstruction by a weighted depth average computed at each pixel x independently:

$$I^{RRCT}(x) = [1 - \alpha(x)]I^{RR}(x) + \alpha(x)I^{CT}(x) \tag{1}$$

where

$$\alpha(x) = \exp\left[\frac{-\beta}{|I^{RR}(x) - I^{CT}(x)|}\right] \tag{2}$$

is an outlier-sensitive regularizer and β is a free variable greater than 0 that can be tuned to adjust the weight given to the CT prior. This formulation gives a higher weight to I^{CT} when the difference between I^{CT} and I^{RR} is high.

3 Results

3.1 Materials

A quantitative validation of our method is performed on cardiac *in silico* phantom data [4,5] consisting of stereo data, CT scan, and ground truth depth. The provided CT data is high resolution: 215 (0.500 mm thick) vertical long axis slices of 512×512 pixels (0.414 mm pixel spacing). The stereo endoscopy data available online are low resolution (360×288 pixels) images captured using a calibrated da Vinci system.

For clinical verification of our method, we fabricated an *ex vivo* phantom using a lamb kidney with artificial tumors. A 16 slice Siemens Somatom CT scanner was used to acquire a high resolution CT volume of the phantom. The resulting volume is composed of 130 (0.600 mm thick) transverse slices of 512×512 pixels (0.215 mm pixel spacing). Stereo endoscopy data was captured with a calibrated da Vinci S system at full HD 1080i resolution. In both studies, we set the value of β to 1 while encoding the CT shape prior into the reconstruction.

3.2 *In Silico* Phantom Study

The three presented reconstruction methods were applied to the dataset. The provided CT was segmented and the resulting shape prior was aligned onto the endoscopic data using the five fiducials visible on the surface of the phantom. The reconstructed depth to surface at each pixel is computed for each of the presented methods and illustrated in Fig. 2, in which the image from the left camera is applied as a texture to enhance visualization. Note how the number of outliers are qualitatively reduced when comparing the results of the reference reconstruction (Fig. 2a) to the robust reconstruction (Fig. 2b). Also note that our proposed method (Fig. 2c) achieves significant qualitative improvements, compared to previous methods, in terms of outliers and smoothness of depth

values. For a quantitative comparison, depth values computed for each pixel of the left image is shown in Fig. 3 alongside respective absolute differences with the provided ground truth distances. Observe how the absolute differences (bottom row of Fig. 3) have been qualitatively reduced from the reference (left) to our proposed method (right). A detailed quantification of theses errors (absolute difference for all pixels) is presented in Fig. 4.

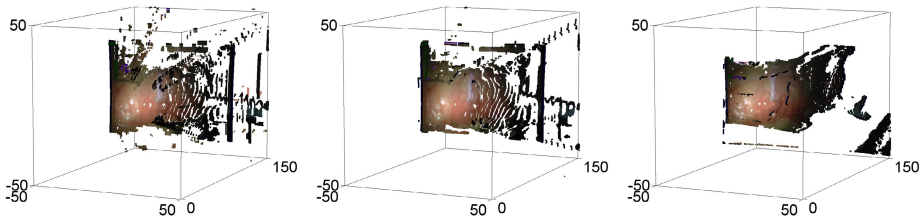


Fig. 2. Reconstructed surface in 3-space (in millimeters) with texture using: (a) simple dense matching method, (b) robust dense matching method, and (c) our proposed method with shape-prior

The results of our proposed method with shape prior in Fig. 4 clearly demonstrate a significant improvement over the two reconstruction methods that do not use a shape prior. Without the shape prior, the mean error (absolute difference between reconstruction and ground truth) was reduced from 17.2 mm to 14.8 mm with the robust algorithm. The median error values remained roughly the same (from 7.7 mm to 7.5 mm). This implies that the few extreme outliers (which had a significant impact on the mean) were removed without significantly changing the results. With our method, the mean and median error values were significantly reduced to 6.1 mm and 4.8 mm respectively. Also note the reduction in variance of error achieved with our method in Fig. 4. The minor increase in minimum error is attributed to the misalignment of the prior.

3.3 *Ex vivo* Lamb Kidney Study

The purpose of this experiment is to qualitatively verify the reduction in mismatched pixels after applying our method to a more realistic *ex vivo* phantom for abdominal RALS. Our phantom provides a good example as it has a smooth texture on its surface, making it very difficult to extract robust depth values by dense matching alone. The CT scan corresponding to the *ex vivo* phantom is segmented in the same fashion as the *in silico* case. The pose is estimated manually from 4 arbitrary landmarks, which were easily identifiable in both CT and stereo images. The three methods were applied to two views of the kidney acquired at different angles; results illustrated in Fig. 5. The reconstruction results of our *ex vivo* phantom exhibit the same pattern of qualitative improvement that

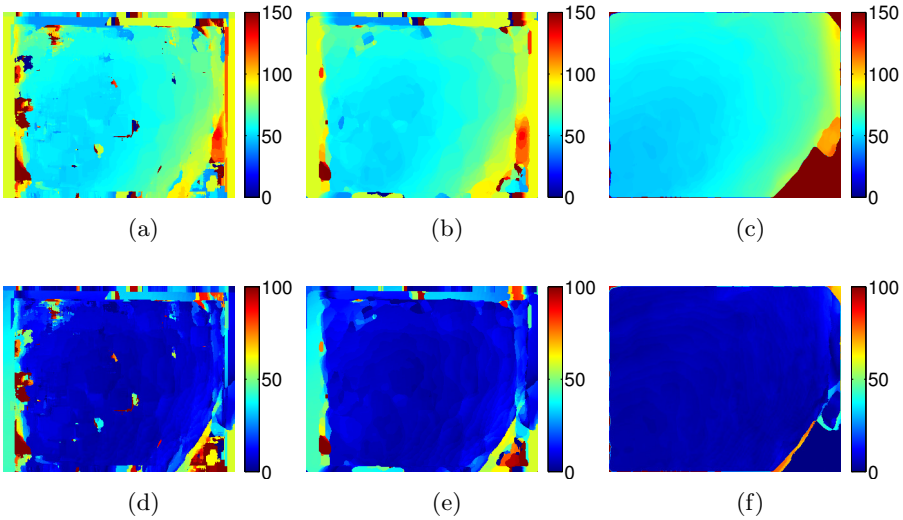


Fig. 3. Top row shows reconstructed distance to surface values (0 mm - 150 mm) using: (a) simple dense matching method, (b) robust dense matching method, and (c) our proposed method with shape prior. Bottom row: (d, e, f) are corresponding absolute difference (0 mm - 100 mm) between the provided ground truth and methods (a, b, c), respectively.

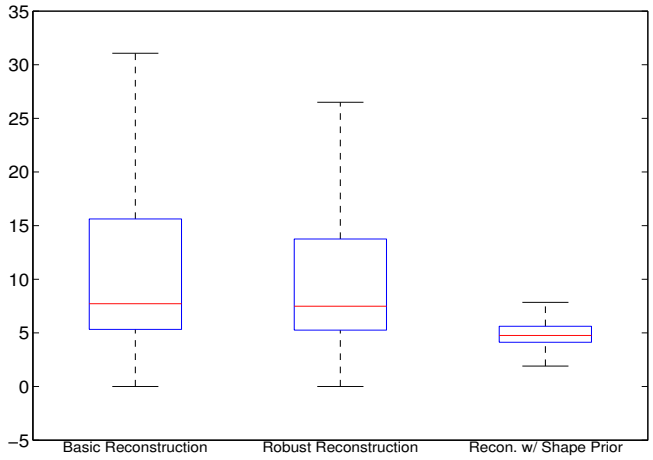


Fig. 4. Box plots of reconstruction errors (absolute difference compared to ground truth) corresponding to each reconstruction method

was quantitatively validated in the *in silico* study. In this particular case, the reconstruction also includes objects other than the kidney, such as the plate. In order to keep these depth values, while incorporating the prior, we have added an extra condition to set $\alpha(x)$ to zero where there is no depth information at $I^{CT}(x)$.

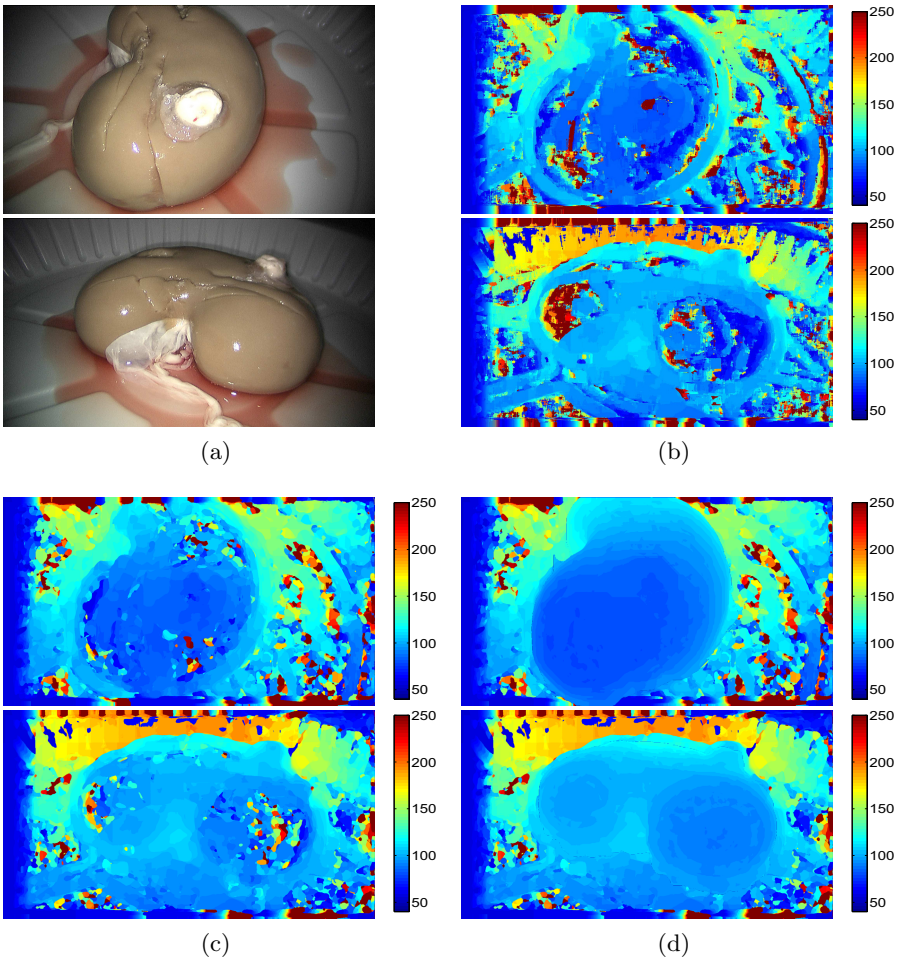


Fig. 5. Reconstructed distance to surface values (45 mm - 250 mm) of our kidney phantom captured at two different angles (a) using: (b) simple dense matching method, (c) robust dense matching method, and (d) our proposed method with shape-prior.

In summary, even without a robust estimation of CT-to-stereo alignment, our method shows a clear visual improvement in areas prone to mismatching (texture-less areas).

4 Conclusion

We have proposed a method that significantly improves stereo surface reconstruction by incorporating priors from pre-operative segmentation into stereo view. The improvement in reconstruction from da Vinci stereo recordings is validated quantitatively on the only publicly available dataset that includes ground truth measurements and we show a 56% reduction in mean absolute difference with ground truth compared to a basic dense matching algorithm. We demonstrate the potential clinical utility of the proposed method in the context of abdominal surgery using a phantom study of an *ex vivo* lamb kidney. Our results show a clear visual improvement in problem areas where mismatches were occurring in methods that don't include a shape prior.

Our proposed prior-informed reconstruction method can be utilized to improve clinical image guidance frameworks that require depth reconstruction such as augmented reality, telestration, feature tracking, dynamic view expansion, biophotonics, and collision detection. For future work, we plan to improve the manual pose alignment stage by adding semi-automatic alignment subroutines. Although our method of fusion accounts for superficial differences and deformations at the surface of the organ, a pose alignment based on similarity transform is not sufficient to account for typical tissue deformations that occur during surgery. In the future, we will investigate the incorporation of a patient-specific biomechanical model of the organ, which after registration can be tracked and non-rigidly deformed to match the current state of the organ. Furthermore, we plan on optimizing the β term with respect to the window size and pursue the option of adaptively varying β based on spatial information. We can tune β to a different value at each pixel depending on NCC score of the best match of I^{RR} at the given pixel.

Acknowledgement. The authors would like to thank Mr. Masoud S. Nosrati for his invaluable input on encoding the shape prior and our clinical collaborators Dr. Abdulla Al Ansari (chairman of surgery at Hammad Medical Corporation) and Dr. Osama Al-Alao for their assistance with the data acquisition. This publication was made possible by NPRP Grant from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

1. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision, 2nd edn. Cambridge Univ. Press (2004)
2. Maier-Hein, L., Mountney, P., Bartoli, A., Elhawary, H., Elson, D., Groch, A., Kolb, A., Rodrigues, M., Sorger, J., Speidel, S., Stoyanov, D.: Optical techniques for 3d surface reconstruction in computer-assisted laparoscopic surgery. Medical Image Analysis 17(8), 974–996 (2013)
3. Guru, K.A., Hussain, A., Chandrasekhar, R., Piacente, P., Bienko, M., Glasgow, M., Underwood, W., Wilding, G., Mohler, J.L., Menon, M., Peabody, J.O.: Current status of robot-assisted surgery in urology: a multi-national survey of 297 urologic surgeons. The Canadian Journal of Urology 16(4), 4736–4741 (2009)

4. Stoyanov, D., Scarzanella, M.V., Pratt, P., Yang, G.-Z.: Real-time stereo reconstruction in robotically assisted minimally invasive surgery. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 275–282. Springer, Heidelberg (2010)
5. Pratt, P., Stoyanov, D., Visentini-Scarzanella, M., Yang, G.-Z.: Dynamic guidance for robotic surgery using image-constrained biomechanical models. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 77–85. Springer, Heidelberg (2010)
6. Pollefeys, M., Koch, R., Van Gool, L.: A simple and efficient rectification method for general motion. In: ICCV, pp. 496–501 (1999)
7. Bernhardt, S., Abi-Nahed, J., Abugharbieh, R.: Robust dense endoscopic stereo reconstruction for minimally invasive surgery. In: Menze, B.H., Langs, G., Lu, L., Montillo, A., Tu, Z., Criminisi, A. (eds.) MCV 2012. LNCS, vol. 7766, pp. 254–262. Springer, Heidelberg (2013)
8. Röhl, S., Bodenstedt, S., Suwelack, S., Kenngott, H., Müller-Stich, B., Dillmann, R., Speidel, S.: Real-time surface reconstruction from stereo endoscopic images for intraoperative registration. In: Proc. SPIE, vol. 7964, p. 796414 (2011)

Multi-atlas and Gaussian Mixture Modeling Based Perirectal Fat Segmentation from CT Images

Soumya Ghose[1], Jim Denham[2], Martin Ebert[3], Angel Kennedy[4], Jhimli Mitra[1], Stephen Rose[1], and Jason Dowling[1]

[1] CSIRO Computational Informatics, 901/16 UQ Health Science Building, Herston, Qld 4029, Australia
[2] School of Medicine and Public Health, University of Newcastle, Callaghan, NSW 2308, Australia
[3] Radiation Oncology, Sir Charles Gairdner Hospital, Hospital Ave, Nedlands WA 6009, Australia and School of Physics, University of Western Australia, 35 Stirling Hwy, Crawley WA 6009, Australia
[4] Radiation Oncology, Sir Charles Gairdner Hospital, Hospital Ave, Nedlands WA 6009, Australia

Abstract. Accurate perirectal fat segmentation in CT images aids in estimating radiation dose delivered to the region of fat around the rectum during radiation therapy treatment of prostate cancer. Such a process is important in determining the resulting toxicity of the neighboring tissues. However automatic or semi-automatic segmentation of the perirectal fat in CT images is a challenging task due to inter patient anatomical variability, contrast variability and imaging artifacts. We propose a combined schema of multi-atlas and multi parametric Gaussian mixture modeling for perirectal fat segmentation in CT images. Multi-atlas based soft segmentation and multi parametric Gaussian mixture modeling aids in identifying the volume of interest (VOI). Thereafter expectation maximization (EM) based soft clustering of the intensities of the VOI refined with positional probabilities of the perirectal fat provides the segmentation of the perirectal fat. The proposed method achieves a mean sensitivity value of 0.88 ± 0.07 and a mean specificity value of 0.998 ± 0.001 with 5 patient datasets in a leave-one-patient-out validation framework. Qualitative results show a good approximation of the perirectal fat volume compared to the ground truth.

Keywords: Multi-atlas, gaussian mixture modeling, computed tomography.

1 Introduction

Prostate cancer is the most commonly diagnosed cancer in Australia. In 2012 more than 18,500 people were diagnosed with prostate cancer and it accounted for 3,235 deaths [5]. Often radiation therapy is used to treat prostate cancer.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 194–202, 2013.

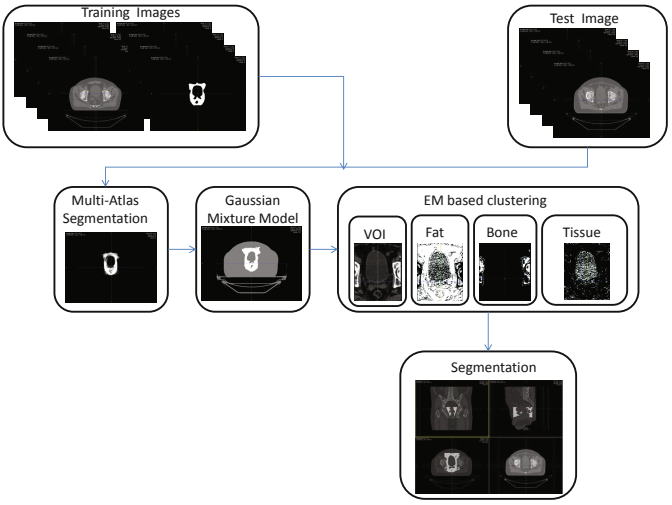


Fig. 1. Schematic representation of our approach

In external beam radiation therapy, high-energy x-ray beams from multiple directions deposit energy (dose) within the tumor to destroy cancer cells. The balancing act in designing a radiation treatment plan is to deliver the maximum dose to a patient's tumor while limiting the amount of radiation given to healthy tissues. Therefore identification of organs and accurate targeting and estimation of dose accumulation in the surrounding tissues of the prostate is absolutely critical.

The randomized androgen deprivation and radiotherapy (RADAR) trial [2] is an ongoing international trial in Australia and New Zealand to determine the effect of androgen deprivation and radiotherapy on prostate cancer patients. The primary objectives of the RADAR trial are to address the following hypotheses 1) that 18 months androgen deprivation in conjunction with radiotherapy is superior to 6 months androgen deprivation prior to and during radiotherapy; 2) that 18 months bisphosphonate therapy will prevent bone loss caused by androgen deprivation therapy and further reduce relapse risk by impeding the development of bony metastases. All patients accrued to the trial received radiotherapy to the prostate with incidental irradiation of surrounding healthy tissues. In the RADAR trial it has been observed that the estimation of radiation dose delivered to the region of fat around the rectum during radiation therapy treatment of prostate cancer is important in determining the resulting toxicity of the neighboring tissues. Fat in CT images has a low electron density and appears dark and visually distinguishable. However perirectal fat is a large and convoluted structure and manual delineation of the 750 datasets of the RADAR trial is difficult and may suffer from inter- and intra-observer variability.

Automatic or semi-automatic computer aided segmentation of the perirectal fat is a challenging task owing to inter-patient shape, size and deformation

variabilities of the bladder, the prostate, and the rectum. Furthermore intensity heterogeneities inside the perirectal fat volume and imaging artifacts adversely affect segmentation accuracies. In the last decade, atlas and statistical shape based prostate segmentation methods in MR images has become popular [6,1,4]. Such methods have produced promising results when validated with large number of datasets. Often machine learning approaches have been adopted for prostate segmentation in CT images. In recent years supervised learning is combined with multi-atlas segmentation for accurate segmentation of the prostate [7]. Motivated by these approaches we propose EM based soft clustering of the intensities of the VOI for segmenting perirectal fat. The VOI is identified in a multi-parametric Gaussian mixture modeling of the intensity and spatial probability distribution obtained in a multi-atlas based segmentation framework. To the best of our knowledge this is the first attempt to segment perirectal fat from CT images. The key contributions of this work are: 1) the use of multi-parametric Gaussian mixture modeling to localize the VOI and 2) use of EM framework to achieve a soft segmentation of the perirectal fat. The remaining paper is organized in the following manner. Section 2 provides a description of the proposed segmentation framework, followed by the results and discussions in Section 3. Finally, the paper is concluded in Section 4.

2 Proposed Segmentation Framework

The proposed method is developed on three major components: multi-atlas based soft segmentation of the perirectal fat, multi-parametric Gaussian mixture modeling to identify the VOI and EM based soft clustering of the voxels in the VOI. The schema of our proposed method is illustrated in Fig. 1.

We present the multi-atlas based soft segmentation of the perirectal fat first followed by multi-parametric Gaussian mixture modeling to localize VOI. The EM based soft clustering of the voxels is presented thereafter.

2.1 Multi-atlas Based Segmentation

We adopt the approach of Klein et al. [6] for our multi-atlas based segmentation procedure. The process involves affine and non-rigid registration of all the training datasets to the test dataset. Closest matching datasets are selected next based on mutual information based similarity and finally the deformed labels of the most similar datasets are fused to achieve a soft segmentation. The process is illustrated in Fig. 2. Multi-atlas based segmentation primarily involves two main steps, registration and label fusion. In the leave-one-patient-out validation framework, in the registration step M(=N-1) training datasets with known labels are registered to the Nth test dataset. Registration is done in two steps, in the first step the training datasets A_i with labels L_i are affine registered to the test dataset P followed by a B-spline [10] based non rigid registration. For affine registration the block matching strategy of Ourselin et al. [9] is adopted. In this method a block of the moving image is compared to the reference image.

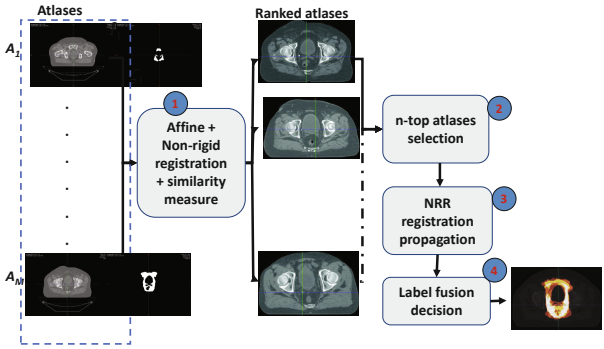


Fig. 2. Schematic representation of multi-atlas based segmentation

The best corresponding block is used to define the displacement vector for affine registration.

Non-rigid registration is performed with B-splines. B-splines consist of a set of control points that can be locally controlled on the image domain. Let $\Omega = \{(x, y, z)|0 \leq x < X, 0 \leq y < Y, 0 \leq z < Z\}$ represent the image domain. The transformation between the moving and fixed images is given by $\mathbf{T}$:$(x, y, z) \mapsto (x', y', z')$, where any point (x, y, z) of the moving image is mapped onto its corresponding point (x', y', z') on the fixed image. Given a mesh of control points on the moving image with a control point defined as $\phi_{i,j,k}$ with uniform spacing of δ mm, the nonrigid transformation $\mathbf{T}$ is defined by B-spline functions as

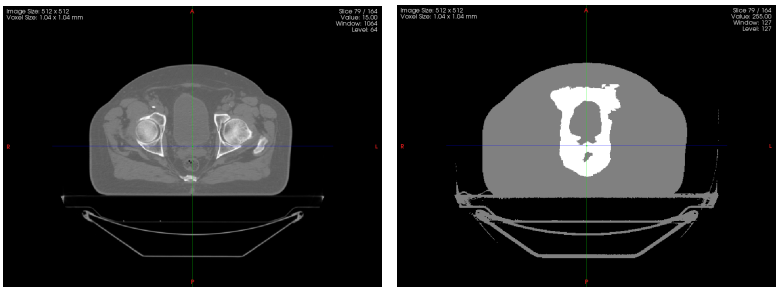
$$\mathbf{T}(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(w) \phi_{i+l,j+m,k+n} \tag{1}$$

where $i = \lfloor x/\delta \rfloor - 1, j = \lfloor y/\delta \rfloor - 1, k = \lfloor z/\delta \rfloor - 1, u = x/\delta - \lfloor x/\delta \rfloor, v = y/\delta - \lfloor y/\delta \rfloor$. $\lfloor . \rfloor$ and $w = z/\delta - \lfloor z/\delta \rfloor$ is the floor function and B_l represents the l^{th} basis function of the cubic B-spline functions such that $B_0(u) = (1 - u^3)/6, B_1(u) = (3u^3 - 6u^2 + 4)/6, B_2(u) = (-3u^3 + 3u^2 + 3u + 1)/6$ and $B_3(u) = u^3/6$.

The resulting coordinate transformation T_i is applied to the training dataset label L_i to produce the resulting label. Normalized mutual information (NMI) [11] as a similarity metric is maximized to achieve the registration. The NMI is an information theoretic measure that tries to reduce the joint entropy of the images and is given by

$$\text{NMI} = \zeta_{similarity} = \frac{H(M) + H(F)}{H(M, F)} \tag{2}$$

where $\zeta_{similarity}$ is the similarity measure for B-splines registration that is maximized in the process, $H(M)$ and $H(F)$ are the marginal entropies of the moving



(a) (b)

Fig. 3. Multi-parametric GMM based clustering framework (a) Axial view of a slice (b) Three class labeling of the same slice

(M) and fixed (F) images respectively, and $H(M, F)$ is the joint entropy of the images. $H(M, F)$ can be written using probability theory as

$$H(M, F) = -\sum_{m,f} p(m, f) \log [p(m, f)], \tag{3}$$

where, $p(m, f)$ is the joint probability distribution of the images obtained from their joint histogram.

The registration stage produces a set of transformed labels $A_i \times T_i$ that must be fused to produce the soft segmentation of the test dataset. The NMI as a similarity metric between the training datasets and the test dataset is chosen for label fusion. In this paper we used majority voting [6] for the fusion of the labels. The choice for labels for fusion is determined from the ratio r_i given by,

$$r_i = \frac{\Phi\left(P, A_i \times T_i\right)}{max_j \Phi\left(P, A_i \times T_i\right)} \tag{4}$$

where $\Phi \in (NMI)$ and max_j is the jth transformed label that produced the highest similarity match. A transformed label was selected if it satisfied $r_i \geq \eta$ where $0 \leq \eta \leq 1$. If $\eta = 0$ all transformed labels were fused and if $\eta = 1$ only the most similar transformed label was considered. The value of η varied between 0.9 to 0.95 for our cases. In the multi-atlas segmentation schema, the Modat et al. [8] approach was adopted for B-spline registration. Multi-resolution affine registration was performed at 3 levels with 50% of the blocks and fixed block size of 4×4×4 voxels. Multi-resolution B-splie based registration was performed at 3 levels with bending energy penalty term 0.005 and control point spacing of 5×5×5 voxels.

2.2 Multi-parametric Gaussian Mixture Model

The soft segmentation obtained from multi-atlas based segmentation in 2.1 aids in identifying the VOI. The results of the multi-atlas based soft segmentation are combined with intensity distribution of the VOI to further refine the results and

localize perirectal fat in a multi-parametric Gaussian mixture model (GMM). We adopt a multi-parametric GMM framework that combines soft classification from multi-atlas based segmentation with GMM of the intensities to minimize the risk of error of localization of the perirectal fat.

To formalize, the probability of a voxel position and intensity being fat is obtained in an EM [3] framework. Given a model X of observed data, a set of latent unobserved data Z and a vector of unknown parameters θ, along with a likelihood function $L(\theta; X, Z)$, the EM algorithm seeks to find the maximum likelihood estimate by iteratively applying the expectation and the maximization steps. In Eq. (5), the expectation step calculates the expected value of the log likelihood function with current estimated parameters θ^t and in (6), the maximization step find the parameters that maximizes this quantity.

$$Q\left(\theta|\theta^t\right) = E_{Z|X,\theta^t}\left[logL\left(\theta; X, Z\right)\right] \tag{5}$$

$$\theta^{t+1} = argmax_\theta\left[Q\left(\theta|\theta^t\right)\right] \tag{6}$$

In our model, the intensity histogram and positional probability obtained from soft segmentation of multi-atlas are approximated with three classes, the perirectal fat, other tissues and the background. Maximum a posteriori estimates of the class means and standard deviations are used to soft cluster the voxels, assigning probabilistic membership values of being in each of the classes as shown in Fig. 3.

2.3 Expectation Maximization Based Soft Clustering

The VOI being identified in 2.2 we adopt a GMM of the intensities of the VOI for soft segmentation of the perirectal fat. We use K-means clustering to roughly cluster the pixels into three classes fat, other tissues and the bones from the intensities. The class means and standard deviations obtained from this rough clustering are then used as the initial estimates in an EM framework to determine the probability of a voxel being perirectal fat (P_{fat}) from intensities. We use eight Gaussian distributions to estimate class probabilities of fat, bone and other tissues to best separate the classes. Class probabilities of the fat, bone and other tissues are extracted depending on the class means and are illustrated in Fig. 4. Affine registration of the training labels aid in determining positional probability of the perirectal fat in the VOI given by P_{pos}. Similarly intensity based voxel probability obtained in EM framework is given as P_{fat}. Maximizing the conditional likelihood of a voxel being perirectal fat ($\widehat{\psi}$) is achieved in a Bayesian framework given by:

$$\widehat{\psi} = argmaxP(\psi|P_{fat}, P_{pos}) \tag{7}$$

The femur head and anal canal were visually identified from CT images for determining the superior and the inferior extent of the fat to localize the zone of the dosimetry plan as advised by the experts. All slices of the segmentation mask above the superior extent and below the inferior extent were discarded.

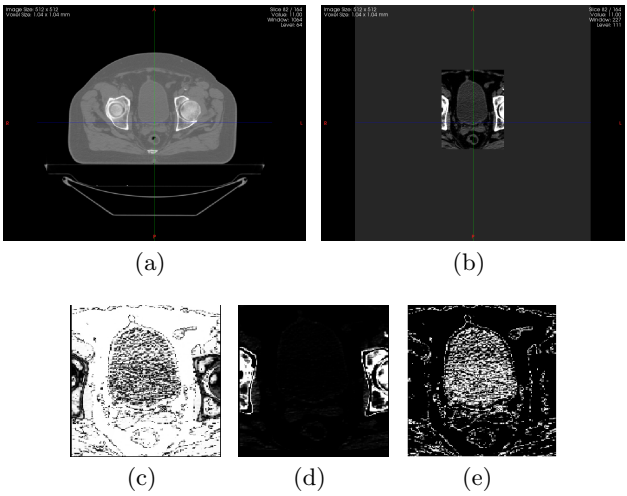


Fig. 4. EM based clustering framework. (a) Axial slice of a dataset (b) VOI determined automatically from 4(a). (c) Perirectal fat probability image (d) Bone probability image (e) Other tissues probability image. Please note that white voxels indicate higher probability.

3 Results

We have validated the accuracy of our approach with 5 anonymised planning CT scans in a leave-one-patient-out validation framework. The image dimensions varied between 512×512×127 to 512×512×164 voxels of size 1×1×2.5 mm. The manual contours were drawn by an experienced radiation oncologist on the Varian Eclipse treatment planning system and transferred via DICOM-RT.

We have used popular segmentation evaluation metrics like sensitivity (Sens.), specificity (Spec.), accuracy (Acc.), Dice similarity coefficient (DSC) and mean surface distance (MSD), to evaluate our method. Sensitivity or true positive rate is given as $Sens. = \frac{TP}{TP+FN}$, specificity or true negative rate is given as $Spec. = \frac{TN}{FP+TN}$, accuracy is given as $Acc. = \frac{TP+TN}{TP+FP+TN+FN}$, where TP=true positive, FP=false positive, TN=true negative and FN=false negative. DSC is a measure of overlap of the same labels (E) between the segmented image ($M(E)$) and the ground truth ($F(E)$) and is given by $DSC = \frac{2(M(E) \cap F(E))}{M(E)+F(E)}$. We present our quantitative results of our approach in Table 1.

In Table 1 we observe a good sensitivity, specificity, accuracy, DSC and MSD values suggesting that we are in good agreement with manual segmentation. The overlap with manual segmentation for some axial slices are illustrated in Fig. 5. A Bayesian model with the least probability of classification error [12] was adopted to build a generalized model for perirectal fat segmentation from a small number of manual segmentation. We adopt the Bayesian estimation in localizing the perrectal fat and in voxel labeling in the VOI minimizing

Table 1. Perirectal fat segmentation quantitative results

Dataset	Sensitivity	Specificity	Accuracy	DSC	MSD (mm)
Set 1	0.81	0.99	0.99	0.78	2.78
Set 2	0.79	0.99	0.99	0.74	2.81
Set 3	0.96	0.99	0.99	0.69	4.48
Set 4	0.94	0.99	0.99	0.76	3.23
Set 5	0.90	0.99	0.99	0.79	2.83
Mean	0.88±0.07	0.99±0.001	0.99±0.001	0.75±0.04	3.23±0.72

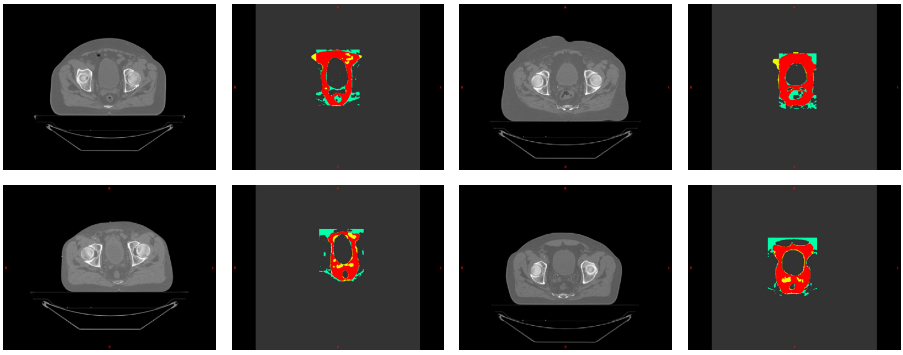


Fig. 5. First and third columns are the axial slices of different datasets, second and fourth columns are the corresponding segmentations achieved by our method. The true positive areas are shown in red, the false negatives are shown in yellow and the false positives are shown in green.

Table 2. Perirectal fat segmentation quantitative results

Dataset	Sensitivity	Specificity	Accuracy	DSC	MSD (mm)
Multi-Atlas	0.67±0.08	0.99±0.02	0.99±0.002	0.59±0.12	4.39±2.29
Multi-atlas and EM	0.88±0.07	0.99±0.001	0.99±0.001	0.75±0.04	3.23±0.72

mis-classification risk at each stage. The effect of incremental learning with multi-atlas and expectation maximization based clustering is demonstrated in Table 2. In Table 2 we observe that the accuracy and the specificity value of multi-atlas segmentation method is very much similar to the values achieved with the combined model of multi-atlas and EM suggesting that multi-atlas successfully identifies the volume of interest and discards true negative. However if we have a closer look at the sensitivity, DSC and MSD values it is evident that inside the volume of interest our combined model of multi-atlas and EM outperforms multi-atlas segmentation. This suggest that the combined model is superior in identifying the true positive inside the volume of interest compared to the multi-atlas schema justifying the use of an incremental learning for solving the problem.

4 Conclusions

A novel schema of multi-atlas segmentation combined with EM based soft clustering with the goal of segmenting the perirectal fat in prostate in 3D CT images has been proposed. The proposed method has shown promising results with a small dataset and is in good consensus with the ground truth. However it is necessary to validate the results with a larger dataset. The results could be improved further with the use of the apriori knowledge of the anatomical landmarks like the femur head, bladder, prostate and the rectum.

References

1. Chandra, S.S., Dowling, J., Shen, K.K., Raniga, P., Pluim, J.P.W., Greer, P.B., Salvado, O., Fripp, J.: Patient specific prostate segmentation in 3D magnetic resonance images. IEEE Trans. Med. Imaging 31(10), 1955–1964 (2012)
2. Clinical Trials, A.G.: Randomised androgen deprivation and radiotherapy (RADAR) trial (2013), `http://www.ClinicalTrials.gov` (accessed on May 19, 2013)
3. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society 39, 1–38 (1977)
4. Gao, Y., Sandhu, R., Fichtinger, G., Tannenbaum, A.R.: A Coupled Global Registration and Segmentation Framework with Application to Magnetic Resonance Prostate Imagery. IEEE Transactions on Medical Imaging 10, 17–81 (2010)
5. Australian Government: Cancer in Australia an overview 2012 (2013), `http://www.aihw.gov.au` (accessed on May 20, 2013)
6. Klein, S., van der Heide, U.A., Lipps, I.M., Vulpen, M.V., Staring, M., Pluim, J.P.W.: Automatic Segmentation of the Prostate in 3D MR Images by Atlas Matching Using Localized Mutual Information. Medical Physics 35, 1407–1417 (2008)
7. Liao, S., Gao, Y., Lian, J., Shen, D.: Sparse patch-based label propagation for accurate prostate localization in CT images. IEEE Trans. Med. Imaging 32(2), 419–434 (2013)
8. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S.: Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine 98(3), 278–284 (2010)
9. Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N.: Reconstructing a 3D structure from serial histological sections. Image Vision Comput. 19(1-2), 25–31 (2001)
10. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: Application to breast MR images. IEEE Trans. Med. Imaging 18(8), 712–721 (1999)
11. Studholme, C., Hill, D.L.J., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition 72(1), 71–86 (1999)
12. Theodoridis, Koutroumbas.: Pattern Recognition, 4th edn. Academic Press (2008)

Selective Search and Sequential Detection for Standard Plane Localization in Ultrasound

Dong Ni[1,⋆], Tianmei Li[1], Xin Yang[1], Jing Qin[2], Shengli Li[3,⋆], Chien-Ting Chin[1], Shuyuan Ouyang[1], Tianfu Wang[1], and Siping Chen[1]

[1]National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, School of Medicine, Shenzhen University, 3688 Nanhai Ave, Shenzhen, P.R.China
[2]Center for Human Computer Interaction, Shenzhen Institute of Advanced Integration Technology, 1068 Xueyuan Ave, Shenzhen, P.R.China
[3]Department of Ultrasound, Affiliated Shenzhen Maternal and Child Healthcare Hospital of Nanfang Medical University, 3012 Fuqiang Rd, Shenzhen, P.R.China
nidong@szu.edu.cn, lishengli63@126.com

Abstract. We present the first automatic solution for localizing fetal abdominal standard plane (FASP) in consecutive 2D ultrasound images. FASP is located in the presence of three key anatomies detected by learning based algorithms, including stomach bubble (SB), umbilical vein (UV), and spine (SP). Traditional detection methods exhaustively scanning the entire image with sliding window algorithms tend not to perform well, since large numbers of regions appear similar to key anatomies. We propose a novel approach by applying local detectors sequentially on the preselected locations of SB, SP and UV. Specifically, we employ segmentation to generate probable locations for SB detection while exploiting a novel accumulative vessel probability algorithm to produce probable locations for SP and UV detection. The sequential scheme automatically excludes detected regions in former steps for subsequent detection, and further limits the search range according to the geometric relationship among anatomies. Experimental results on 100 fetal abdomen videos show that our method significantly outperforms traditional methods that only use local detector.

Keywords: Ultrasound, standard plane, anatomy detection, AdaBoost, selective search, sequential detection, vessel probability map.

1 Introduction

The acquisition of standard planes is crucial for accurate biometric measurements and diagnosis during medical ultrasound (US) imaging. However, acquiring standard plane images can be time-consuming, skill- and experience-dependent, and lacking in consistency and reproducibility. It is often cited as one of the inherent disadvantages of US compared to other imaging modalities such as CT and MRI. Recently, a number of automatic methods for detecting standard planes from 3D US [1–3] have been proposed to increase productivity and/or decrease inter- and intra-observer variability. Although 3D US is a

⋆ Corresponding author.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 203–211, 2013.

promising imaging modality for prenatal examination, currently and in the near future 2D exams still dominate the field due to the better imaging quality, the wider availability of 2D scanners and the preference of experienced users [4]. Zhang et al. [5] proposed to select the standard plane of gestational sac based on cascade AdaBoost classifier and local context information. Rahmatullah et al. [6] proposed to detect abdomen anatomical landmarks using AdaBoost. Liu et al. [7] proposed to identify an acceptable fetal head scan plane by employing an advanced Active Appearance Model.

Above-mentioned classifier based methods commonly applied local detectors on the entire image or volume with an exhaustive scan algorithm. The detection accuracy can potentially be degraded since large numbers of regions appear similar to key anatomies. For example, in fetal abdominal US, shadows, abdominal aorta (AO), inferior vena cava (IVC) and hypoechoic spinal cord are often detected as SB or UV. Rahmatullah et al. [8] proposed to limit the search range of the local detector only to the locations preselected by global features, and the performance of this method was thus improved. However, experimental results show that similar looking regions still lead to suboptimal detection accuracy.

In this paper, we propose a novel selective search and sequential detection approach to automatically localize the fetal abdominal standard plane (FASP) in consecutive 2D US images. FASP is located when SB, SP and UV are all detected in one frame [9]. In our design, the clinician sweeps the scan plane through the fetal abdomen in conventional manner, however, the task of selecting FASP is performed automatically. To the best of our knowledge, it is the first automatic solution for localizing FASP in 2D US.

The overall framework of our proposed method is as follows. First, the region containing abdomen (namely region of interest, ROI) is detected using its AdaBoost classifier to reduce the search range. Next, the specific and efficient selective strategy for each anatomy is designed to generate probable locations for detection. Specifically, we propose to use segmentation to automatically generate the probable locations of SB, and propose to obtain the probable locations of UV and Spine (SP) by using a novel accumulative vessel probability approach. Thus a local detector is applied only to these probable locations. Such selective search strategies can significantly improve the detection accuracy by pre-rejecting similar looking regions. Furthermore, preliminary experiments show that among three target anatomies, SB can be most easily detected using a trained AdaBoost classifier, and regions nearby SP are often detected as UV mistakenly. Therefore, we employ a sequential strategy to detect SB, SP and UV sequentially, which can improve the detection accuracy by automatically excluding detected regions in former steps for subsequent detection, and limiting the search range according to the geometric relationship among anatomies. Finally, FASP is located from the consecutive US images in the presence of SB, SP and UV in one frame.

2 Classifier Training

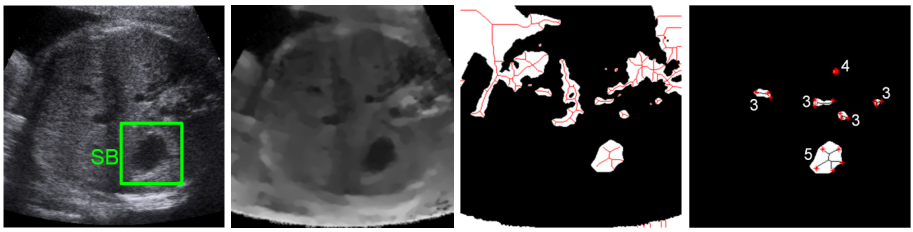
We exploit the AdaBoost learning algorithm [10] to select a set of Haar-like features from given training samples, and train the classifier. AdaBoost has been

proved to be an efficient algorithm and widely used in object detection.In our study, four classifiers were trained each for detecting ROI, SB, SP and UV. Positive samples were generated by cropping image regions that contain the anatomical objects while negative samples were extracted randomly from the background and some of them have an overlap of 20% to 40% with a positive sample. Note that the cropped image region was normalized into a square of 80×80 pixels.

3 Selective Search and Sequential Detection

3.1 Selective Search for SB Detection by Automatic Segmentation

SB is visualized as a dark circular structure in fetal US. We propose to use segmentation as a selective search strategy to generate a set of probable locations and hence improve the robustness of its detection. This procedure is shown in Fig. 1. The compensation method for acoustic attenuation [11] is first employed on the input ROI image (Fig. 1 (a)) to remove shadows and then the widely used speckle reducing anisotropic diffusion (SRAD) algorithm [12] is applied to reduce speckle noises. The filtered image (Fig. 1 (b)) is further segmented by a fully automatic and efficient OTSU thresholding algorithm [13] with 256 histogram bins. It is observed that there still exist a number of false candidates after segmentation (Fig. 1 (c)). Compared with the circular structure of SB, these regions usually have irregular appearance or small number of pixels. Our statistical analysis on a set of 200 training images further proves that the number of skeleton endpoints of most false regions is significantly different from the true SB region. Therefore we propose to further eliminate false candidates by counting the number of skeleton endpoints of each segmented region. The candidates are removed if the number is lower than the minimum threshold (3 in our experiments) or greater than the maximum threshold (8 in our experiments). Fig. 1 (d) shows the final probable locations of SB. The digits on the image are the numbers of skeleton endpoints of segmented regions. The lines are skeletons of segmented regions. Finally, detection of SB is carried out by applying its AdaBoost classifier on the sliding windows only centered on the candidate skeleton points.

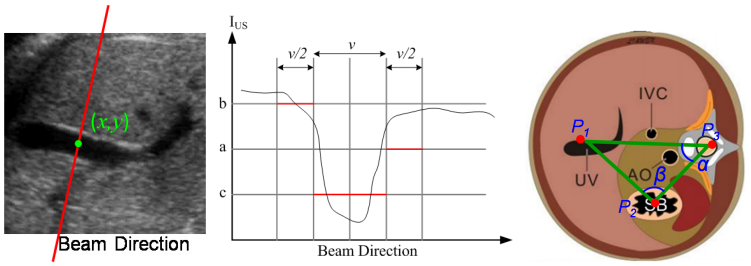


(a) (b) (c) (d)

Fig. 1. (a) ROI, (b) Result after shadow removal and filtering, (c) Segmentation result and skeleton, (d) Preselected SB locations

3.2 Selective Search for SP and UV Detection by AVPM

The UV is visualized as a dark cresent shape in fetal US. The SP is near the AO and characterized as the hypoechoic spinal cord surrounded by hyperechoic spines. Therefore it is reasonable to predict locations of SP by localizing AO and the hypoechoic spinal cord, which look similar to vessels. In this study, we propose to generate probable locations of UV and SP by detecting salient vessel features based on the vessel probability map (VPM) algorithm [14]. However, the original VPM algorithm does not perform well in UV detection due to the relative low contrast between UV and its surrounding tissues along the acoustic direction. In this regard, we propose a novel accumulative VPM (AVPM) method to enhance the vessel feature by summing up the VPM of each frame in the video.



(a) (b) (c)

Fig. 2. (a), (b) Illustration of dip calculation, (c) Geometric relationship among anatomies

First the VPM is obtained from each frame by calculating the vessel probabilities (VP) of all pixels. For each pixel in the image, we compute a dip value which quantifies its likeness to the typical vessel appearance of a dark centre with a relative bright rim. Then this pixel's VP is obtained from a 2D look-up table using its dip value and intensity value as indices. The VP look-up table P_{vessel} is pre-generated from a set of 100 annotated fetal abdomen images using Eq. 1. The UV, AO and hypoechoic spinal cord are manually segmented to generate the annotation image.

$$P_{vessel}(I, I_{dip}) = \frac{\text{Number of vessel pixels with intensity } I \text{ and dip value } I_{dip}}{\text{Number of US pixels with intensity } I \text{ and dip value } I_{dip}} \quad (1)$$

The dip value is calculated by a novel modification of the original method, in order to enhance the vessel feature along the beam direction. The dip value is calculated along two directions (along and normal to the beam direction) instead of only along the beam direction. First the original image is transformed from Cartesian to polar coordinates along the beam direction. Then, as shown in Fig. 2 (a) and (b), the dip value $D(x, y)$ of a pixel at position (x, y) is defined thusly: for each possible vessel size v, compute the dip value $d = \min(a - c, b - c)$, where c is the mean intensity value along (or normal to) the beam direction within the linear range of v centered at (x, y), and a, b are the mean intensity values along (or normal to) the beam direction within a range of $v/2$ just distal

(right), or proximal (left), respectively, to the hypothetical vessel lumen. In our experiments, the vessel size v is from 8 to 16 pixels. The maximal value of d for all values of v is defined as $D_i(x, y)$ where $i = 1$ for along, and $i = 2$ for normal to the beam direction. $D(x, y)$ is then simply defined as $D_1 + D_2$, which efficiently capture the likelihood of (x, y) being a central point of a vessel in any direction relative to the US beam.

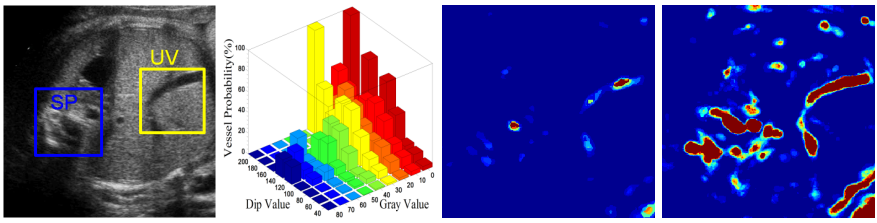


Fig. 3. (a) ROI, (b) 2D look-up table, (c) VPM, (d) AVPM

Fig. 3 shows a typical US image (Fig. 3 (a)) at a number of stages from the original image to the AVPM. Fig. 3 (b) shows the 2D look-up table. High dip value and low intensity indicate the high probability of vessel. The result shows that the VPM (Fig. 3 (c)) generated from one frame is not good enough to indicate the vessel probability. Compared with the VPM, the AVPM shown in Fig. 3 (d) indicates the better probable locations of UV and SP.

AVPM is further thresholded to generate the probable locations of SP and UV. Then SP is detected using its classifier along the selective locations. At last, UV is detected by its classifier along the left pre-selective locations, which are further constrained by the geometric relationship among UV, SB, and SP. One of the authors, an experienced radiologist, proposed a restriction of probable locations by geometric angles, which was further validated by our experiments. As illustrated in Fig. 2 (c), point P_1, P_2 and P_3 are the centers of UV, SB and SP regions, respectively. α and β are the angles between line P_3P_1 and line P_3P_2, line P_2P_1 and line P_2P_3, respectively, which should be smaller than a threshold according to the anatomical relationship. Based on the experiments on 500 training images, we set $\alpha < 80°$ and $\beta < 120°$. Therefore probable locations of UV are eliminated if this requirement is not satisfied.

4 Experimental Results and Discussion

We used 1664 expert annotated fetal US abdomen images for generating the training samples shown in Table 1. Some of them were also used for training parameters in our proposed methods. Besides, 100 fetal abdomen videos acquired from 100 pregnant women were used for detection test. Fetal gestational age was between 20 and 36 weeks. All images and videos were acquired using a Siemens Acuson Sequoia 512 US scanner with a 4-6 MHZ transabdominal probe from Shenzhen Maternal and Child Healthcare Hospital. Conventional US sweep was

performed to obtain the videos on pregnant women in the supine position. Each sweep lasted approximately 2-6 s and each video contains 17-48 frames. The FASP in each video was manually selected by a radiologist with more than five years of experience in obstetrics US, as ground truth for comparison with the results obtained by our proposed method.

Table 1. Details of training data

	Positive	Negative		Positive	Negative
ROI	1140	1940	SP	640	1800
SB	620	1000	UV	430	600

We experimented with three different detection methods: local detector based on AdaBoost algorithm, selective search guided detection (SSD) and selective search guided sequential detection (SSSD). For the AdaBoost method, the image was exhaustively scanned at multiple scales using sliding windows. In the SSD method, local detectors were applied to the preselected probable locations without considering the detection sequence. In these experiments, a ground-truth object is first considered being detected if 75% of its area is covered by detected boxes, and further confirmed by clinicians.

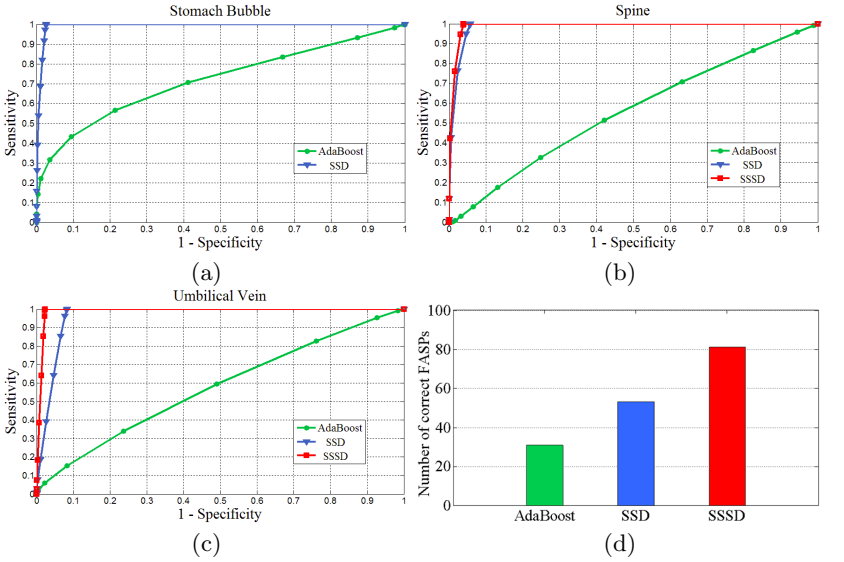
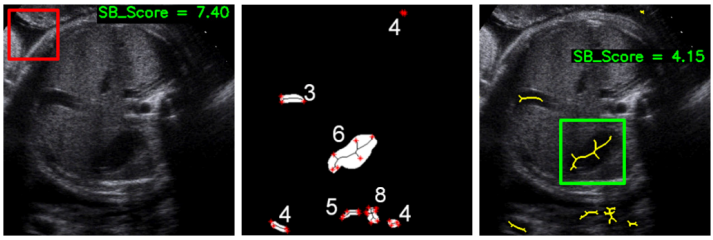


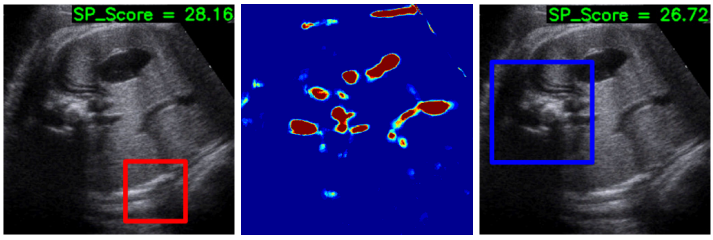
Fig. 4. (a) ROC plot for Detection of SB, (b) ROC plot for Detection of SP, (c) ROC plot for Detection of UV, (d) Comparison of FASP localization using three methods

We first compared the performance of three methods on detection of anatomical objects using ROC curves, as shown in Fig. 4 (a)-(c). It is observed that the detection accuracy of both SSD and SSSD are significantly better than that of AdaBoost and the proposed SSSD method achieves the best results. Specifically,

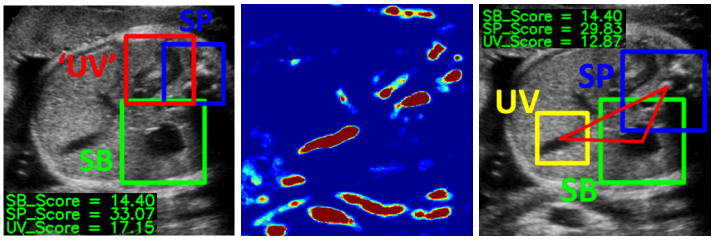
the UV detection accuracy of the SSSD method is greatly improved due to the use of sequential detection. Some examples shown in Fig. 5 further demonstrate that the proposed method can eliminate many false positives by selective search and sequential detection.



(a) Left: false positive result of SB detected by AdaBoost (red box). Middle: preselected SB locations based on segmentation. Right: false positive result is corrected with true positive (green box) by preselected locations (yellow lines).



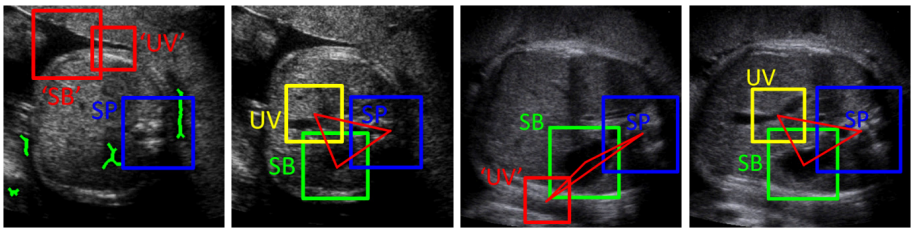
(b) Left: false positive result of SP detected by AdaBoost (red box). Middle: AVPM result for SP detection. Right: false positive result is corrected with true positive (blue box) by AVPM.



(c) Left: false positive result of UV detected by AdaBoost (red box). Middle: AVPM result for UV detection. Right: false positive result is corrected with true positive (yellow box) by AVPM and sequential detection.

Fig. 5. Comparison of detection results using our method and AdaBoost

Finally, we compared the performance of three methods for the localization of FASP from the 100 videos. As shown in Fig. 4 (d), we can correctly locate FASP in 31, 53 and 81 videos using AdaBoost, SSD and SSSD method respectively from the 100 videos, demonstrating the effectiveness of our method. Fig. 6 illustrates two examples of FASPs located by AdaBoost method and the proposed



(a) (b) (c) (d)

Fig. 6. (a) and (c) are false FASPs located from two videos by AdaBoost method, while (b) and (d) are correct FASPs located from the same videos of (a) and (c) by the proposed method

method. In Fig. 6 (a), a false FASP is located by false positives of SB and UV obtained by AdaBoost method, while from the same video, the true FASP (Fig. 6 (b)) can be located by true positives of SB and UV detected by our method. In Fig. 6 (c), another false FASP is located due to the false positive of UV obtained by AdaBoost method while in our method the the false positive of UV can be corrected by sequential detection and geometric relationship among anatomical objects, and hence the correct FASP is obtained. The proposed system is implemented by the mixed programming technology of Matlab and C++. It takes several minutes to localize the FASP from one video using a 3GHz Intel Pentium CPU and 3G RAM-based workstation.

5 Conclusion

This paper presents the first automatic solution for localizing FASP from consecutive 2D US images. We propose a novel selective search and sequential detection approach to detect key anatomies for FASP localization in fetal abdomen US images. Segmentation method and a novel AVPM method are used to generate the probable locations of SB, SP and UV. The geometric relationship among anatomies is further used to eliminate false candidates obtained by AVPM. Experimental results demonstrate our proposed method significantly outperforms the traditional AdaBoost method. The selective search and sequential detection mechanisms are general, and can be potentially adopted to the detection of other anatomies in medical images. Our future work will focus on the automatic measurement of abdomen circumference in fetal US based on the detected FASPs.

Acknowledgements. The work was supported in part by National Science Foundation of China (Project No. 61101026, No. 81270707 and No. 61233012), in part by Shenzhen key Basic Research Project (Project No. 201101013, No. JCYJ20130329105033277 and No. JSE201109150013A).

References

1. Lu, X., Georgescu, B., Zheng, Y., Otsuki, J., Comaniciu, D.: Autompr: Automatic detection of standard planes in 3d echocardiography. In: 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2008, pp. 1279–1282. IEEE (2008)
2. Carneiro, G., Amat, F., Georgescu, B., Good, S., Comaniciu, D.: Semantic-based indexing of fetal anatomies from 3-d ultrasound data using global/semi-local context and sequential sampling. In: CVPR 2008, pp. 1–8 (2008)
3. Rahmatullah, B., Papageorghiou, A., Noble, J.A.: Automated selection of standardized planes from ultrasound volume. In: Suzuki, K., Wang, F., Shen, D., Yan, P. (eds.) MLMI 2011. LNCS, vol. 7009, pp. 35–42. Springer, Heidelberg (2011)
4. Elliott, S.T.: Volume ultrasound: the next big thing? British Journal of Radiology 81(96), 8–9 (2008)
5. Zhang, L., Chen, S., Chin, C.T., Wang, T., Li, S.: Intelligent scanning: Automated standard plane selection and biometric measurement of early gestational sac in routine ultrasound examination. Medical Physics 39(8), 5015–5027 (2012)
6. Rahmatullah, B., Sarris, I., Papageorghiou, A., Noble, J.: Quality control of fetal ultrasound images: Detection of abdomen anatomical landmarks using adaboost. In: ISBI 2011, pp. 6–9 (2011)
7. Liu, X., Annangi, P., et al.: Learning-based scan plane identification from fetal head ultrasound images. SPIE Conference Series, p. 8320 (2012)
8. Rahmatullah, B., Papageorghiou, A., Noble, J.: Integration of local and global features for anatomical object detection in ultrasound. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 402–409. Springer, Heidelberg (2012)
9. Dudley, N.J., Chapman, E.: The importance of quality management in fetal measurement. Ultrasound in Obstetrics and Gynecology 19(2), 190–196 (2002)
10. Viola, P., Jones, M.J.: Robust real-time face detection. International Journal of Computer Vision 57, 137–154 (2004)
11. Girard, M.J.A., Strouthidis, N.G., et al.: Shadow removal and contrast enhancement in optical coherence tomography images of the human optic nerve head. Investigative Ophthalmology & Visual Science 52(10), 7738–7748 (2011)
12. Yu, Y., Acton, S.: Speckle reducing anisotropic diffusion. IEEE Transactions on Image Processing 11(11), 1260–1270 (2002)
13. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics 9(1), 62–66 (1979)
14. Penney, G., Blackall, J., Hamady, M., et al.: Registration of freehand 3d ultrasound and magnetic resonance liver images. Medical Image Analysis 8(1), 81–91 (2004)

Rib Detection in 3D MRI Using Dynamic Programming Based on Vesselness and Ridgeness

Yolanda H. Noorda, Lambertus W. Bartels, Max A. Viergever, and Josien P.W. Pluim

Image Sciences Institute, University Medical Center Utrecht
Heidelberglaan 100, 3584 CX Utrecht, The Netherlands
Y.H.Noorda@umcutrecht.nl

Abstract. In this paper, a fully automatic method is proposed to detect the ribs in 3D MRI. The purpose of the detection is MR-guided HIFU treatment of liver lesions, in which the ribs should be avoided. Rib segmentations are required for treatment planning and they may also be used for motion tracking during treatment. The rib detection results can serve as an initialization to automatic rib cage segmentation. The algorithm is based on surface detection and dynamic programming. First, the outer surface of the rib cage is detected. Vesselness and ridgeness are computed to highlight elongated structures. The ribs are tracked simultaneously on a 2D projection of the vesselness in the surface, using dynamic programming. Finally, the extracted lines are backprojected into the original 3D volume. Preliminary results of this algorithm are presented on data of five subjects. The results were evaluated by visual inspection of the backprojected lines in 3D. It was checked whether a line belonged to the correct rib and whether it stayed inside this rib. Overall, our algorithm was capable of detecting the ribs that were visible in the images. Testing on five volunteers yielded one failure. The remaining four results were satisfactory. Our method seems suitable to serve as initialization to a full rib cage segmentation in MRI.

Keywords: Ribs, segmentation, detection, MRI, vesselness, ridgeness, dynamic programming, image-guided therapy.

1 Introduction

Non-interventional treatments are becoming increasingly popular. For treatment of liver lesions, ablation by MR-guided High Intensity Focused Ultrasound (MR-HIFU) is a promising technique. During HIFU-therapy, an ultrasound beam is used to create a high intensity focus at the tumor site in order to heat the tissue, which eventually leads to cell death. The temperature is monitored by MRI.

However, several problems need to be solved before HIFU can be used in practice for liver tumor treatment. Due to liver motion, it is difficult to maintain a focus of sufficient intensity at the tumor. In addition, heating of the ribs

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 212–220, 2013.

should be avoided. Bone is known to have a high absorption coefficient; therefore, ultrasound waves traveling through the bone can cause serious pain for the patient. This poses even more limitations on the ultrasound beam configuration [1]. To allow accurate treatment planning, it would be highly favorable to obtain the position of the ribs automatically. The ultrasound beam can then be placed in such a way that the waves will travel through the intercostal space, while still being able to create a sufficiently high temperature at the focus. For this purpose, an accurate segmentation of the rib cage in MRI is required.

The rib cage consists of bone and cartilage. The first seven ribs are attached to the sternum by the costal cartilage. The 8^{th}, 9^{th} and 10^{th} ribs join with the costal cartilage of the 7^{th} rib. The posterior part of the ribs consists of bone. The floating ribs are not attached to the costal cartilage and consist of bone only.

The cartilage appears bright in the scans, but the edge contrast is low. Since bone gives no signal on MRI, the bone itself is only visible by means of the contrast with the surrounding tissue. The transition of cartilage to bone is hardly visible in MRI. This makes the segmentation task particularly challenging.

Rib cage segmentation in the literature is mainly performed on CT [2,3] or radiographs [4,5]. Since the bones are well visible in CT images, this problem is different from detecting the ribs in MRI. This means that the same task on MRI requires a different approach. Segmentation of the rib cage in MRI has not been described in the literature before. There are examples of segmentation of bony structures in MRI, which use prior knowledge and deformable models [6,7]. Segmentation of cartilage is mainly done on knee cartilage [8,9]. However, this is not comparable with our segmentation problem.

In this paper, we propose a fully automatic detection of the rib cage in MRI. This can be used as an initialization for rib cage segmentation during treatment planning, to determine the optimal ultrasound beam configuration. In addition, such a full rib cage segmentation can be used during treatment to update the position of the ribs with respect to the ultrasound focal spot.

2 Materials and Methods

2.1 Data

Since HIFU therapy of the liver is not used in the clinic yet, clinical data of the rib cage on MRI were not available. Therefore, MRI scans of the rib cage were obtained on five healthy volunteers. It is assumed that the results of the segmentation on real patients will not differ from our results on volunteers, since the pathology will be in the liver and not in the rib cage.

The volunteers were placed in prone position, to simulate HIFU treatment. A 3D T1-weighted gradient echo coronal MRI scan (Philips Achieva, TE/TR 1.96/4.032 ms, flip angle 10°, in-plane resolution 1.372×1.372 mm^2, slice thickness 2 mm) was acquired. The scan consisted of 90 slices and was acquired in breathhold.

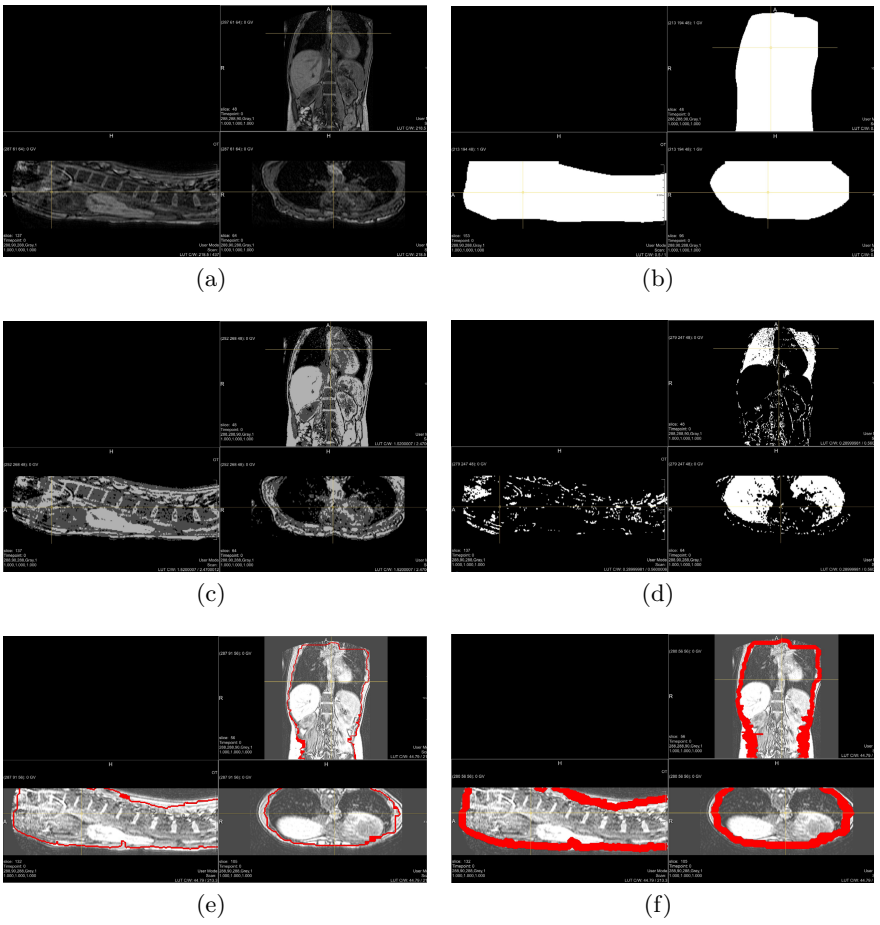


Fig. 1. Finding the surface enclosing the rib cage. (a)The original volume. (b) The body is separated from the background. (c) The clustering method yields three classes. (d) The voxels belonging to the first cluster are taken. (e) The contour around the rib cage. (f) The final surface with a thickness of 11 voxels.

2.2 Method Overview

The rib detection method consists of three main steps. First, the surface that encloses the rib cage is found. In the second step, a 3D vesselness filter is applied on the voxels in this surface to highlight elongated structures. Then, the surface is unfolded to obtain a 2D image, on which the corresponding vesselness values are projected. On this image, ridgeness is calculated to highlight the centerlines of all elongated structures. The ridgeness values are used to calculate a cost image, on which dynamic programming is performed to find the first seven ribs. Finally, these paths are mapped back into the 3D volume to obtain the centerlines of the ribs. Below, each step will be explained in more detail.

2.3 Surface of the Rib Cage

The body is separated from the background by a combination of morphological operations. By taking the largest connected component, only the torso remains (Fig. 1(b)). A fuzzy clustering algorithm is used on the original volume to roughly separate the different structures in the body: cluster 1 contains dark structures like lungs, vessels and bone, cluster 2 contains structures with an intermediate intensity like muscle and organs, and cluster 3 contains bright structures like fat and the skin (Fig. 1(c)). We take only those voxels that belong to cluster 1. We remove the voxels that belong to the body-background border by erosion of the body with a $3 \times 3 \times 3$ spherical kernel, to get rid of possible misclassified skin or fat voxels. A binary mask containing the lungs, vessels and bones is left (Fig. 1(d)). From this mask, we take the outer boundaries. Then we perform morphological closing in 3D. This yields a volume of interest of which the outer boundary encloses the rib cage. A level set filter is applied to obtain a closed contour (Fig. 1(e)). We take this contour with a margin of 5 voxels inwards and outwards as the final surface. The margins yield a surface thickness of 11 voxels (Fig. 1(f)), to make sure that the surface contains the rib cage.

2.4 Unfolding the Surface

A 3D vesselness filter is applied for bright and dark structures separately, using the filter proposed by Frangi et al.[10]. The vesselness value of the voxels belonging to the 3D surface is stored. The surface is then unfolded to obtain a 2D projection. By using the 2D projection, one degree of freedom can be discarded, such that it is ensured that the paths searched for will follow the contour of the rib cage. In addition, it enables significantly faster computation of the optimal paths compared to a 3D search.

The surface is projected in 2D by radial sampling of the transverse slices. Since the transverse shape of the body is not exactly circular, but rather elliptical, an ellipse is fitted to the central transverse slice. This ellipse serves as a reference shape for every slice. This way, the sampling points are distributed equally over the surface. The 3D surface is resampled on this ellipse for every slice, by increasing the angle of 0° to 360° in 2000 steps. Since the 3D surface is not a real surface but still has a thickness of 11 voxels, the voxel with the highest vesselness output is taken for every angle and its 3D location is stored. The vesselness projection is shown in Fig. 2(a).

2.5 Location of the Ribs

The 2D projection of the vesselness also includes the lungs, since the posterior part of the rib cage is not included in the scan. To avoid distortion of the paths by the lungs, they are excluded automatically by iterative thresholding and connected component analysis. Then the 2D image is cropped.

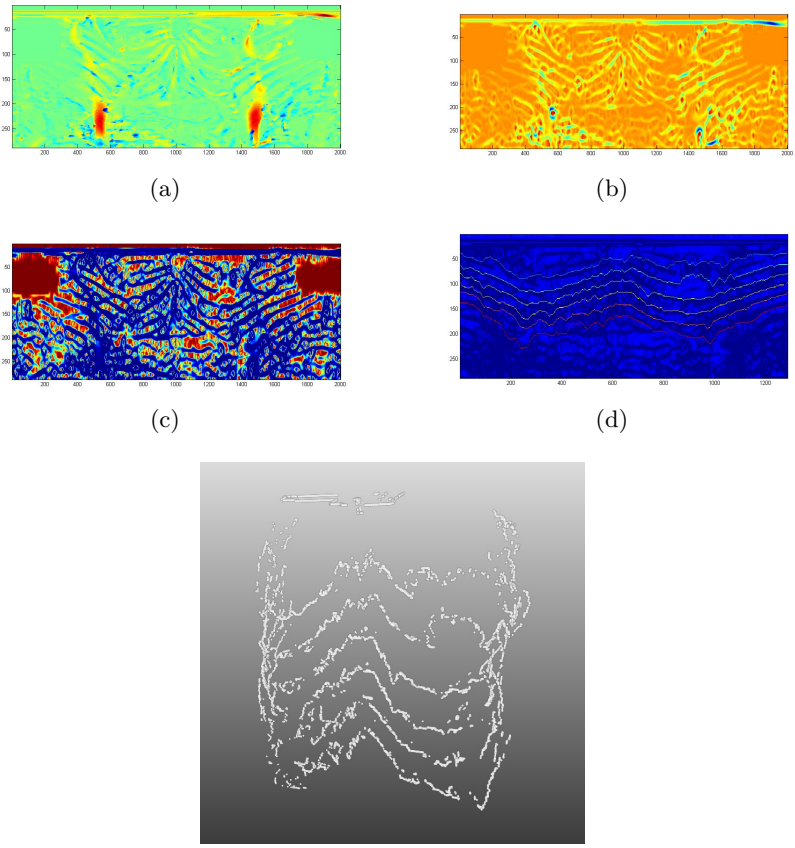


Fig. 2. Method overview. (a) The projected vesselness. (b) A ridgeness filter is applied. (c) The cost image. (d) Seven minimal paths are found. (e) The rendering in 3D.

The elongated structures in 3D are shown as lines in the 2D projection. To obtain the center of these lines, a 2D ridge filter is applied (Fig. 2(b)). This filter calculates the main principal curvature as follows:

$$L = \frac{t^{\gamma}}{2}\left(L_{xx} + L_{yy} - \sqrt{\left(L_{xx} - L_{yy}\right)^2 + 4L_{xy}^2}\right) , \tag{1}$$

where L_{xx} and L_{yy} are the second derivative of L in the x- and y-directions, respectively, and γ is set to $\frac{3}{4}$. As ridges are points where the principal curvature reaches a maximum, the output of this filter will be high for the ribs.

A dynamic programming algorithm is used to track the ribs in the resulting image. Our purpose is to detect the first seven ribs. Therefore, seven minimal paths are searched through this image simultaneously.

The cost function depends on the ridgeness of the current pixel. The algorithm searches for a minimal path; therefore, the exponent of the negative of the ridgeness value is taken to ensure a value close to zero for high ridgeness and a value close to one for no ridgeness (Fig. 2(c)). The cost function at node n_i is therefore defined as:

$$C(n_i) = \exp\left(\frac{-L(n_i)}{a}\right) , \tag{2}$$

where a is set to 0.01 to yield a steep descent. With this function, seven cumulative cost images are built on which seven minimal paths need to be searched. Every pixel in a cost image represents the accumulated cost associated with traveling from the starting point to that specific point. Since the paths should have a minimal distance from each other, we are not looking for the exact minimal paths in the cost images. Instead, we search for a structure of seven paths over these seven images, with a certain distance between them, such that the total cost of this structure is minimal. Therefore, we will perform backtracking from seven end points. At each step, we penalize paths that approach each other. This implies that at every node, we add a distance penalty to the accumulated cost at the neighboring nodes, and then take the neighboring node with the lowest cost. The distance penalty for a path p_k at column position i is defined as:

$$D(p_k, i) = \sum_{\substack{j\in[1,N]\\ j\neq k}} \frac{\sigma}{D_E\left(n_{k,i}, n_{j,i+1}\right)} * \exp\left(\frac{D_E\left(n_{k,i}, n_{j,i+1}\right)}{2\sigma^2}\right) . \tag{3}$$

Here, D_E is the Euclidean distance between two nodes, N is the total number of paths and σ is set to 10 voxels, which approximates the average radius of a rib. Since the unfolded surface is reasonably symmetric, each final path should start and end at the same height. These heights are not known beforehand. Therefore, the algorithm runs twice. First we look for seven minimal paths without putting any constraints on start and end points. The heights at which these initial paths end, indicate the height of the ribs. To guarantee closed contours, the algorithm runs a second time, but now we force the paths to start and end at these specific heights. Therefore, this time the cost image is not built as usual, but for every start point a separate graph structure is created. In this graph, only nodes that lead back to the same height as the start position are connected. This finally yields seven closed contours which represent the ribs (Fig. 2(d)).

2.6 3D Centerlines of the Ribs

By using the 3D coordinates that were stored during the unfolding of the surface, the 3D centerlines of the ribs can be restored. This is shown in Fig. 2(e).

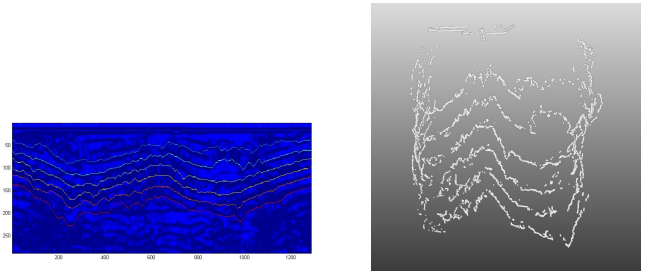
3 Results

Results are shown as lines on the 2D projections, as well as 3D renderings. In general, all ribs that were visible in the 2D projections were found. When the

surface missed a part of the rib cage, this affected the cost image. The algorithm is able to overcome gaps in the structures; however, if too many gaps occur, it is easy for paths to jump to another rib, thereby distorting the neighboring paths. In Table 1, the results are evaluated by counting the number of correctly detected ribs and evaluating the structure of the rib cage in terms of how well the lines are located inside the ribs. The symbol + was assigned to cases in which the detected ribs had a correct shape, such that the ribs were parallel to each other and no major 3D distortions occurred. A − was assigned to cases in which parts of the ribs were missed or major 3D distortions occurred, such as moving the wrong direction due to wrong surface initialization or jumping between different ribs.

For the third volunteer, the surface missed a large part of the rib cage. As a consequence, the lower part of the rib cage was not contained in the 2D projection. For the other images, some paths were incorrect, because part of them belonged to another rib.

Figure 3 shows a typical output. It can be seen that the rendering is less accurate at the transition points between bone and cartilage. This was expected, since this transition is almost invisible in the images. Although the lines are not perfectly smooth, the algorithm is able to detect the shape of the ribs. Obviously, these are very preliminary results. However, it seems that they could serve well as an initialization to a full segmentation.



(a) (b)

Fig. 3. A typical result

Table 1. Evaluation of the output

Volunteer	Number of correct ribs	Correct localization
1	7	+
2	6	+
3	3	-
4	5	-
5	4	+

4 Discussion

We have shown that our algorithm is capable of finding the first seven ribs in abdominal MR images.

The robustness of our algorithm depends mainly on finding a surface that encloses the rib cage. When this surface was accurate, the projection of the ribs on 2D was also correct. The thickness of 10 voxels of the surface provided a margin for misalignment of the rib cage. Problems occurred mainly when the surface missed a part of the ribs. This caused either gaps in the 2D projection, or distortions from non-rib structures in the surface. Gaps of a few voxels could be overcome by the algorithm. Distortions caused false disconnected outputs in 3D. Therefore, it might be feasible to add a final step that filters disconnected pixels in the 3D output. The most important step is thus to ensure that all rib voxels are contained in the volume.

The elliptical sampling ensured that the sampling points were equidistant. By performing radial sampling, such as simply using polar coordinates, areas of the surface with high curvature (such as the lateral rib cage) are sampled less than areas with lower curvature. Elliptical sampling therefore yielded a more accurate 2D projection on which rib structures were more connected. To determine the eccentricity of the elliptical sampling, an ellipse was fitted to the body. We therefore chose the axial slice that contained the center of mass as the slice on which the ellipse was fitted. Although the shape of the body is not perfectly elliptical or similar in all axial slices, this slice was found to represent the axial shape of the body sufficiently. The elliptical sampling did not introduce new errors, since it only had an effect on stretching parts with high curvature.

To calculate the ridgeness, principal curvature was chosen as a measure, because it assigns a ridgeness value to every voxel, as opposed to a binary filter. This enabled detection of strong and weak ridges simultaneously. This way, a smooth cost function could be obtained, to allow for a robust optimal paths search.

The cost function that was used to track the ribs ensures that we find seven minimal paths that do not intersect each other. The parameter σ, that indicates how large this distance should be, was set to a value of 10 voxels, which approaches the radius of an average rib. Overall, our method was able to detect the rib cage in 3D MRI. The first seven ribs could be tracked. The other ribs are not connected to the sternum, but to the cartilage of the 7^{th} rib or to the spine. They could therefore not be detected with our algorithm, but this could be added as an extension in the future.

The dynamic programming approach guaranteed that closed contours were found. Since all seven ribs were searched for simultaneously, paths could influence each other. Therefore, ribs that were more difficult to find could still be detected.

The results are preliminary, but they are promising and could be used as an initialization to a fully automatic segmentation scheme.

Acknowledgments. This work was financially supported by the project Mediate (Patient friendly medical intervention) in the framework of the EU research programme ITEA (Information Technology for European Advancement).

References

1. Quesson, B., Merle, M., Köhler, M.O., Mougenot, C., Roujol, S., De Senneville, B.D., Moonen, C.T.: A Method for MRI Guidance of Intercostal High Intensity Focused Ultrasound Ablation in the Liver. Med. Phys. 37, 2533–2540 (2010)
2. 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)
3. 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)
4. van Ginneken, B., ter Haar Romeny, B.: Automatic Segmentation of Lung Fields in Chest Radiographs. Med. Phys. 27, 2445–2455 (2000)
5. Xu, X.-W., Doi, K.: Image Feature Analysis for Computer-Aided Diagnosis: Accurate Determination of Ribcage Boundary in Chest Radiographs. Med. Phys. 22, 617–626 (1995)
6. Lorigo, L.M., Faugeras, O., Grimson, W.E.L., Keriven, R., Kikinis, R.: Segmentation of Bone in Clinical Knee MRI Using Texture-Based Geodesic Active Contours. In: Wells, W.M., Colchester, A., Delp, S. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 1195–1204. Springer, Heidelberg (1998)
7. Schmid, J., Magnenat-Thalmann, N.: MRI Bone Segmentation Using Deformable Models and Shape Priors. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 119–126. Springer, Heidelberg (2008)
8. 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)
9. 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)
10. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale Vessel Enhancement Filtering. In: Wells, W.M., Colchester, A., Delp, S. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)

Modeling and Simulation of Soft Tissue Deformation

Yuping Duan[1], Weimin Huang[1], Huibin Chang[2], Wenyu Chen[1], Kyaw Kyar Toe[1], Jiayin Zhou[1], Tao Yang[1], Jiang Liu[1], Soo Kng Teo[3], Chi Wan Lim[3], Yi Su[3], Chee Kong Chui[4], and Stephen Chang[5]

[1] Institute for Infocomm Research, A*STAR, 1 Fusionopolis Way, Singapore, 138632
[2] Department of Mathematical Sciences, Tianjin Normal University, 241 Weijin Rd, Tianjin, China, 300387
[3] Institute of High Performance Computing, A*STAR, 1 Fusionopolis Way, Singapore, 138632
[4] Department of Mechanical Engineering, National University of Singapore, 21 Lower Kent Ridge Rd, Singapore, 119077
[5] National University Hospital, 5 Lower Kent Ridge Rd, Singapore, 119074

Abstract. A stable and accurate deformable model to simulate the deformation of soft tissues is a challenging area of research. This paper describes a soft tissue simulation method that can deform multiple organs synchronously and interact with virtual surgical instruments accurately. The model we used in our method is a multi-organ system by point masses and springs. The organs that anatomically connect to each other are jointed together by high stiffness springs. Here we propose a volume preserved mass-spring model for simulation of soft organ deformation. It does not rely on any direct constraint on the volume of tetrahedrons, but rather two constraints on the length of springs and the third constraint on the direction of springs. To provide reliable interaction between the soft tissues and kinematic instruments we incorporate the position-based attachment to accurately move the soft tissue with the tools. Experiments have been designed for evaluation of our method on porcine organs. Using a pair of freshly harvested porcine liver and gallbladder, the real organ deformation is CT scanned as ground truth for evaluation. Compared to the porcine model, our model achieves a mean absolute error 1.5024 mm on landmarks with a overall surface error 1.2905 mm for a small deformation (the deformation of the hanging point is 49.1091 mm) and a mean absolute error 2.9317 mm on landmarks with a overall surface error 2.6400 mm for a large deformation (the deformation of the hanging point is 83.1376 mm). The change of volume for the two deformations are limited to 0.22% and 0.59%, respectively. Finally, we show that the proposed model is able to simulate the large deformation of the liver and gallbladder system in real-time calculations.

Keywords: Physically based modeling, mass-spring, time integration, surgery simulation, volume preservation.

1 Introduction

With the development of laparoscopic techniques, surgery simulation becomes an increasingly relevant alternative to traditional training methods. Elastic deformation models have been greatly studied in the last two decades for the surgery simulation. Nealen et al. [1] gave a good overview of the deformable models used in computer graphics. Generally speaking, the vast number of techniques used for soft tissue deformation

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 221–230, 2013.

modeling can be classified into two different categories: the heuristic models such as Mass-Spring Model (MSM) and continuum-mechanical approach such as the Finite Element Method (FEM). The FEMs are based on a physical model of deformation and able to accurately compute complex deformations of soft tissues. Compared to a linear FEM, which is numerically fast but not well suited to moderate deformations and rotations, a nonlinear FEM has advantage that it is more reliable to rotations and large deformations. However, it lead to a requirement of both high computational cost and large memory usage in simulation. To achieve real-time deformation, non-linear FEMs rely on either the pre-computation or GPU-based accelerations. On the other hand, MSMs, which are robust to topology changes and large deformations, are also widely used to model deformable objets [2] due to the ability in generating dynamic behaviors in real time.

A MSM is a discrete model, which consists of a set of point masses connected by ideal weightless elastic springs. To model a solid 3D object, a tetrahedral mesh modeling both the surface and internal structure of the object can be constructed using the MSM. Suppose a MSM is composed of n point masses $\boldsymbol{x}_i \in \mathbb{R}^3$, $i = 1, \ldots, n$ with mass $m_i \in \mathbb{R}$ and the forces acting on the point is $\boldsymbol{f}_i \in \mathbb{R}^3$. The geometric state of all points is simply $\boldsymbol{x} \in \mathbb{R}^{3n}$, $\boldsymbol{f} \in \mathbb{R}^{3n}$ and $\boldsymbol{m} \in \mathbb{R}^{3n \times 3n}$, respectively. The relation between the acceleration and force can be described by Newton's Law of motion as follows:

$$\boldsymbol{m}\ddot{\boldsymbol{x}} = \boldsymbol{f}, \tag{1}$$

where $\ddot{\boldsymbol{x}}$ is the second derivative of the position with respect to time.

In this work, we present a soft tissue simulator that uses a fast tetrahedral mass spring model to calculate soft tissue deformation, where the model parameters are selected according to soft tissue properties. The position based interaction with kinematic virtual surgical tools is applied to achieve accurate attachment. The deformation of our multi-organ system is realized according to (1) by accumulating external and internal forces on point masses. We model the connecting tissues between soft organs as springs with high stiffness, the so-called repulsive springs. With the repulsive springs, no collision detection is required for the synchronous deformation illustrated by the liver and gallbladder. Volume preservation is important for realistic modeling of soft yet solid tissues. Lasseter [3] states, "The most important rule to squash and stretch is that, no matter how squashed or stretched out a particular object gets, its volume remains constant." To reach this point, two constraints on the length of springs and the third constraint on the direction of springs are constructed to serve as post processing to the MSM. The constraints introduce extra non-linearity to the conventional MSM. Unlike prior works [4,5], in which the volume of tetrahedrons is investigated in computation, our constraints act on mass points instead of tetrahedrons. The proposed volume preserved MSM is validated on the real deformation of a porcine liver with gallbladder. The CT scanned deformation is compared with the computational deformation of MSM. The comparison results demonstrate the accurate performance of our method. Finally, we achieve a real-time dynamic system with reasonable accuracy of organ deformation and interaction with a kinematic virtual tool for a simplified mesh model.

2 Volume Preserved MSM

The deformation is estimated based on Verlet integration according to total forces. We discuss different forces in our system, which contain both the external forces (pulling attachment and gravitation) and the internal forces (spring forces, damping forces and contact forces). Novel constraints on point masses are incorporated into the system to realize the volume conservation for soft tissues.

2.1 Deformation Estimation

For the numerical simulation, we first separate Equation (1) into two coupled first order equations by introducing the velocity $\boldsymbol{v} \in \mathbb{R}^3$ as follows

$$\begin{cases} \dot{\boldsymbol{v}} = \boldsymbol{f}(t)/\boldsymbol{m}, \\ \dot{\boldsymbol{x}} = \boldsymbol{v}. \end{cases} \tag{2}$$

We use Verlet integration, which is among the simplest and most popular explicit schemes in real-time applications to solve (2). The basic idea is to keep the position at previous time $t - \Delta t$ and use this information to obtain a more accurate prediction for $t + \Delta t$. Taylor expansion of the position in the two time directions yields

$$\begin{aligned} \boldsymbol{x}(t + \Delta t) &= \boldsymbol{x}(t) + \dot{\boldsymbol{x}}(t)\Delta t + \frac{1}{2}\ddot{\boldsymbol{x}}(t)\Delta t^2 + \frac{1}{6}\dddot{\boldsymbol{x}}(t)\Delta t^3 + O(\Delta t^4), \\ \boldsymbol{x}(t - \Delta t) &= \boldsymbol{x}(t) - \dot{\boldsymbol{x}}(t)\Delta t + \frac{1}{2}\ddot{\boldsymbol{x}}(t)\Delta t^2 - \frac{1}{6}\dddot{\boldsymbol{x}}(t)\Delta t^3 + O(\Delta t^4). \end{aligned}$$

By adding the above two equations and bringing in (1) and ignoring the high order terms, we have the so-called Verlet integration scheme for the MSM as follows

$$\begin{cases} \boldsymbol{x}(t + \Delta t) = \boldsymbol{x}(t) + \boldsymbol{v}(t)\Delta t + \boldsymbol{f}(t)\Delta t^2/\boldsymbol{m}, \\ \boldsymbol{v}(t + \Delta t) = (\boldsymbol{x}(t + \Delta t) - \boldsymbol{x}(t))/\Delta t. \end{cases} \tag{3}$$

2.2 Forces Modeling

Attachment. In order to accurately move vertices of soft tissues along with the interacted kinematic instruments, we use the position-based attachment [6]. The position of selected vertices are updated at every time step to coincide with the motion of the kinematic instrument. Suppose the initial and objective position of mass point $\boldsymbol{x}_i$ are $P_0(\boldsymbol{x}_i)$ and $P(\boldsymbol{x}_i)$ and the attachment is done in n iteration, the movement of the point $\boldsymbol{x}_i$ in each iteration is $(P(\boldsymbol{x}_i) - P_0(\boldsymbol{x}_i))/n$.

Gravitation. The force of gravity is acting on every point mass in the system and written as $\boldsymbol{f}^g(\boldsymbol{x}_i) = m_i\boldsymbol{g}$, where $\boldsymbol{g}$ is the gravitational acceleration.

Spring Forces. Springs are modeled with linear elasticity. The force acting on mass i generated by the spring connecting i and j is in direct proportion with the extension of the spring. Therefore, according to Hooke's Law, the spring force is defined as follows

$$\boldsymbol{f}_i^s = k_{(i,j)}(\|\boldsymbol{x}_j - \boldsymbol{x}_i\| - l_0) \cdot \frac{\boldsymbol{x}_j - \boldsymbol{x}_i}{\|\boldsymbol{x}_j - \boldsymbol{x}_i\|}, \tag{4}$$

where $k_{(i,j)}$ is the spring stiffness and l_0 is the rest length of spring (i, j).

Damping Forces. Due to imperfect elasticity of physical bodies, energy dissipation occurs during the deformation. We use spring damping to represent the viscous force. These damping forces are defined as

$$\boldsymbol{f}_i^d = d_{(i,j)} \frac{(\boldsymbol{v}_j - \boldsymbol{v}_i) \cdot (\boldsymbol{x}_j - \boldsymbol{x}_i)}{\|\boldsymbol{x}_j - \boldsymbol{x}_i\|} \cdot \frac{\boldsymbol{x}_j - \boldsymbol{x}_i}{\|\boldsymbol{x}_j - \boldsymbol{x}_i\|}, \tag{5}$$

where $d_{(i,j)}$ is the spring's damping constant of spring (i, j).

Contact Forces The gallbladder is connected to the lower surface of the liver at the gallbladder fossa by connecting tissues. Therefore, when external forces are applied to the liver/gallbladder, there exist contact forces in the contact surface of the liver and gallbladder. We generate repulsive springs in the contact surface of the liver and gallbladder to model contact forces. The contact forces are the forces combined of spring forces and damping forces from repulsive springs.

2.3 Deformable Constraints

We obtain both the position $\boldsymbol{x}$ and velocity $\boldsymbol{v}$ after Verlet integration. Next, we introduce constraints on the position as the post-processing process to regularize the simulation results. These constraints introduce extra non-linearity to the conventional MSM.

Spring Length Correction. It is well-known when a concentration of large forces occurs in a small region of soft tissues, the simulation result of a MSM falls into the problem of local deformation ("super-elastic" effects) [7]. We design a pair of constraints on the spring length based on the deformation rate, which is defined as $\tau = l - l_0/l_0$. To identify the influence of the stretch and compression in the deformation, different deformation rates are used for the over-stretching compensation and over-compression compensation.

Over-stretching Compensation. We set a critical stretching rate τ_s to the springs to protect the spring from being stretched too much. More specifically, when the length of the spring exceeds $(1 + \tau_s) \times l_0$, the constraint is applied to try to push the spring back to $(1 + \tau_s) \times l_0$. Therefore, we define the over-stretching correction as the following inequality constraint:

$$C_{stretch}(\boldsymbol{x}_i, \boldsymbol{x}_j) = (1 + \tau_s) \times l_0 - \|\boldsymbol{x}_i - \boldsymbol{x}_j\| \geq 0, \tag{6}$$

If the above inequality constraint is not satisfied, we compute the corrections on the point $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ along the gradient of $C_{stretch}$. The formulae of the correction term $\Delta\boldsymbol{x}_i$ and $\Delta\boldsymbol{x}_j$ are given as follows:

$$\begin{aligned} \Delta\boldsymbol{x}_i &= \frac{1}{2} \times C_{stretch}(\boldsymbol{x}_i, \boldsymbol{x}_j) \cdot \frac{\boldsymbol{x}_i - \boldsymbol{x}_j}{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}, \\ \Delta\boldsymbol{x}_j &= -\frac{1}{2} \times C_{stretch}(\boldsymbol{x}_i, \boldsymbol{x}_j) \cdot \frac{\boldsymbol{x}_i - \boldsymbol{x}_j}{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|}. \end{aligned}$$

After the implementation of over-stretching constraint, $\forall \boldsymbol{x}_i$, $i = 1, \ldots, n$, we sum up all the corrections $\Delta\boldsymbol{x}_i$ contributed by the edges containing the point $\boldsymbol{x}_i$, namely

$$\boldsymbol{x}_i = \boldsymbol{x}_i + \frac{1}{m} \sum_{\mathcal{E}_i} \Delta\boldsymbol{x}_i, \tag{7}$$

where $\mathcal{E}_i$ denotes the set of edges containing the point $\boldsymbol{x}_i$ and m is the total number of edges in $\mathcal{E}_i$.

Over-compressing Compensation. On the other hand, as long as the length of the spring is less than $(1-\tau_c)\times l_0$, τ_c is the critical compressing rate, we use another constraint to push the spring back to $(1-\tau_c)\times l_0$. The constraint for over-compressing springs is defined as

$$C_{compress}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \|\boldsymbol{x}_i - \boldsymbol{x}_j\| - (1-\tau_c)\times l_0 \geq 0, \tag{8}$$

The method to compute the update from (8) and the following direction constraint is similar to over-stretching constraint by computing the gradient of the constraint.

Spring Direction Correction. During the simulation, sudden change of the spring direction may cause serious problem, such as instability and self-collision. Therefore, we define another constraint on the spring direction to guarantee that the direction of the spring is within a critical rotation angle θ in certain iterations. More specifically, we define the constraint on the spring direction as follows

$$C_{direction}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \theta - \arccos\left(\frac{(\boldsymbol{x}_j - \boldsymbol{x}_i)\cdot(\boldsymbol{x}_j^p - \boldsymbol{x}_i^p)}{\|\boldsymbol{x}_j - \boldsymbol{x}_i\|\|\boldsymbol{x}_j^p - \boldsymbol{x}_i^p\|}\right) \geq 0, \tag{9}$$

where $\boldsymbol{x}_i^p$ denote the previous state of the system, whose initial value is the initial position of meshes and is updated every certain number of iterations to the current position.

3 Validation and Results

We use a newly harvested porcine liver with controlled deformation by an external hanging thread to measure the simulation performance. The organs in small and large deformation are CT scanned and segmented as the ground truth for the evaluation. The real deformation is generated by pulling the liver using a stick hanged by a cotton thread to certain positions as in Fig. 1. Before we deform the liver and gallbladder, we put 10 markers on the liver surface in Fig. 1(a). In the experiments, we track the CT scanned position of these makers before and after the deformation and compare them with estimated results.

(a) Rest status (b) Small deformation (c) Large deformation

Fig. 1. Experiments setup. From left to right is CT scan I, CT scan II and CT scan III.

3.1 Experimental Setup

Similar to [8], we define four types of point masses in our MSM as follows: 1) The force that deforms the liver is generated by pulling a small stick underneath the liver Lobe 1 through a cotton thread (Fig. 1(b)). To mimic the pulling force in the simulation, ***attachment points*** along the stick and the thread are manually selected and the position of these points are described based on the tool position in the experiment. 2) We observe that there is nearly no deformation of Marker 8-10 and the other lobe (Lobe 2) for both deformations. For simplicity, all point masses in the posterior and Lobe 2 of the liver that touch to the ground surface are regarded as ***fixed points*** during the simulation. 3) To generate the repulsive springs between the liver and gallbladder, we select the points on the gallbladder that are in contact with the liver and then find the ***joint points*** on the liver by the smallest distance. 4) All other liver and gallbladder points are ***free points***, the movement of which are completely determined by the resulting forces acting on them.

Point Masses. Assume the mass of a tetrahedron is equally divided among its vertices. The mass m_i of mass point i is estimated as:

$$m_i = \sum_{\forall j \in \mathcal{T}_i} \frac{1}{4} \rho V_j, \tag{10}$$

where $\mathcal{T}_i$ is the union of all tetrahedrons that contain point i, V_j is the volume of tetrahedron j and ρ is the tissue density. In our experiments, the mass density of the liver is 1060 kg/m^3 [9] and the mass density of the bile (gallbladder) is 1000 kg/m^3 [10].

Spring Stiffness. The parametrization of spring stiffness provides certain stress-stain relationship of soft tissues. In [11], a formula was established to compute spring stiffness for a regular tetrahedron with unique spring length based on an isotropic elastic material with Young's modulus E. For our irregular tetrahedral formulation, we calculate an equivalent edge length l_e from the volume for each tetrahedron element e, i.e., $l_e = (V_e \frac{12}{\sqrt{2}})^{\frac{1}{3}}$. According to [11], we compute spring stiffness for body springs from:

$$k_{(i,j)} = \sum_{e \in \mathcal{T}_{(i,j)}} \frac{2\sqrt{2}}{25} l_e E, \tag{11}$$

where $\mathcal{T}_{(i,j)}$ is the set of tetrahedrons that contain the edge (i, j). For our experiments, the Young's modulus for the liver and gallbladder are $E = 3.5$ kPa [12] and $E = 1.5$ kPa [10], respectively. On the other hand, we use very stiff spring parameters for repulsive springs due to their function in preventing collision.

Spring Damping. In [2], the authors have given the formula to calculate the damping constant to ensure the best behavior consistency for different and combined resolution. For the spring connecting point mass m_i and m_j with initial length l_0, we use the following formula to compute the damping constant:

$$d_{(i,j)} = \frac{2\sqrt{k_{(i,j)}(m_i + m_j)}}{l_0}. \tag{12}$$

3.2 Evaluation Criteria

To compare the simulation results with the porcine model (PM), the absolute error of the markers in estimated meshes by the MSM is computed as follows:

$$\epsilon_i = \|\boldsymbol{x}_i^{PM} - \boldsymbol{x}_i^{MSM}\|. \tag{13}$$

In (13), $\boldsymbol{x}_i$ denotes the position of the ith node.

3.3 Model Evaluation

The tetrahedral mesh models of the liver and gallbladder are generated in *TetGen* environment. We use the tetrahedral mesh containing 12958 tetrahedrons to represent the liver and the tetrahedral mesh containing 5423 tetrahedrons to represent the gallbladder. We use the obtained MSM to estimate the deformations of the porcine liver with gallbladder. The parameters are selected as $\tau_s = 0$, $\tau_c = 0$, $\theta = \pi/3$ and the $\boldsymbol{x}^p$ is updated by every 10 iterations in the spring direction constraint. The deformation occurring at the attachment points of the small and large deformation is the same for the porcine model and computational model, which is 49.1091 mm and 83.1376 mm, respectively. In Table 1, we record 1) the distance of the rest status and different deformations on CT scans, 2) the estimated distance of the initial mesh and simulated meshes of the small and large deformation and 3) the absolute error ϵ of the position on CT scans and the estimated position for each marker. Besides, we also examine the interior deformation by selecting two more landmarks on the vessels inside the liver. The mean absolute errors for the internal markers are 1.6245 mm and 2.6431 mm for the small and large deformation, which are similar in the error ranges. The simulation results are shown in superposed images of initial status in Fig. 2. We observe that the total volume of the liver and gallbladder has changed 0.22% and 0.59% after simulation for small and large deformation, respectively.

Table 1. Performance analysis of the obtained MSM with estimated distance and absolute error. The unit of the distance is mm.

Marker Index	Small deformation			Large deformation		
	CT distance	Estimated distance	ϵ_i	CT distance	Estimated distance	ϵ_i
1	28.7457	31.2330	2.7084	58.4756	61.6608	4.4024
2	14.7709	15.7711	1.5620	38.9963	41.4500	3.4310
3	10.7932	13.0685	2.3248	30.8646	27.5830	3.5275
4	10.6665	11.7714	1.1251	28.8311	31.7520	3.0398
5	4.9477	5.7381	1.2013	19.3340	17.8716	3.9076
6	4.3042	5.3377	1.1951	6.6086	6.7473	1.6022
7	1.8389	1.8769	0.4002	2.3040	2.6638	0.6117
Mean Error	–	–	1.5024	–	–	2.9317

In addition, we measure the overall performance of our model in approximating the porcine deformable model by selecting 200 points on the top surface of the liver where deformation obviously happens. For the two deformations, the overall mean error is 1.2905 mm and 2.6400 mm, respectively. The box plot of the error is shown in Fig. 3 (a). We render the scene using OpenGL and refresh rate of 15 frames per second (FPS) can be achieved on PC with CPU Intel(R) Core(TM) i7-2600 running at 3.40GHz.

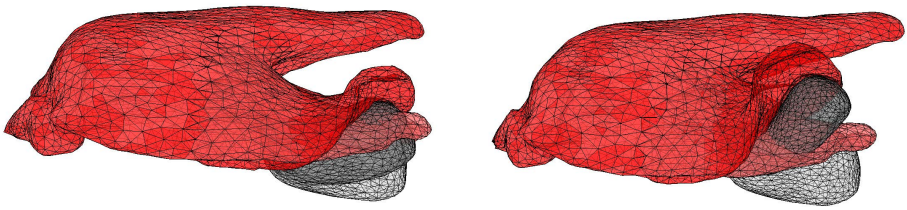


Fig. 2. Simulation results of the small and large deformation shown in superposed image of the mesh without deformation. In the figures, the red and black meshes represent the deformed liver and gallbladder while the pink and white meshes represent the rest liver and gallbladder.

Based on the above evaluation of the obtained MSM, we can make the following observations:

1. The comparison between the simulation results and porcine liver deformation shows that our model can achieve the simulation with good accuracy.
2. The results demonstrate that our system can efficiently handle large deformations and is capable of multi-organ simulation. The deformation is well distributed in the system without any local deformation effect.
3. By well setting the deformation rates (i.e., $\tau_s = 0$ and $\tau_c = 0$), the simulator can model the incompressibility in deformable solids by enforcing volume preservation over point masses instead of individual tetrahedra. The idea is similar to [13] to avoid computation of volume for each tetrahedron, but our solution is from the discrete point of view instead of FEM.

Remarks: Noted that the precision of the marker position selected in CT images is limited by the CT scan resolution, which is 0.488 mm $\times$0.488 mm $\times$1 mm, the CT artifacts and subjective differences by human beings. The mean difference of the marker position, that is the system error, is around 0.5 pixel, corresponding to 0.3099 mm.

3.4 Different Mesh Resolution

The deformation accuracy is mesh-density related. With sparser mesh resolution, the accuracy will drop. For the liver mesh with 1698 tetrahedrons and the gallbladder mesh with 1157 tetrahedrons, the mean error of markers is given by 2.5879 mm and 5.2715 mm and the change of volume is 0.56% and 1.24% for the small and large deformation, respectively. Similarly, we select 100 points in the same region and examine the overall mean error of the two deformations, which is 2.7279 mm and 4.2015 mm. The error is

slightly increased compared to the dense mesh model. The box plot is given in Fig. 3 (b). For the sparse mesh resolution experiment, the refresh rate of 40 FPS was achieved on the same PC environment.

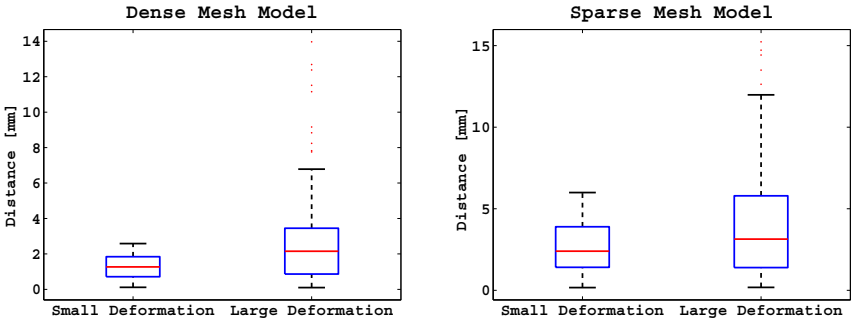


Fig. 3. Performance of our volume preserved MSM for deformation

Based on the test on mesh resolutions, we make another observation about the proposed method: Since all constraints in the obtained model act on springs, the computational costs can be greatly reduced when the mesh resolution decreases. With relatively sparse mesh models, our volume preserved MSM can obtain real-time calculations for surgical simulations.

4 Conclusion

We presented an explicit formulation of a MSM with novel constraints to simulate the deformation of soft tissues. In our approach, we modeled the contact forces between the liver and gallbladder by creating repulsive springs with high stiffness in the contact surface. We benefit from modeling the solid deformable tissues by enforcing volume preservation over point masses instead of individual tetrahedron. By position based attachment, the model can reliably and accurately track the tool movement. It was evaluated by comparing the simulation results with the real porcine liver and gallbladder deformation. The estimated results show that the proposed model can perform similar deformation as the porcine model and can be realized in real-time calculation for surgical simulation.

Acknowledgements. This work is supported by a research grant 102-1480009 from BEP Office, A*Star, Singapore.

References

1. Nealen, A., Müller, M., Keiser, R., Boxerman, E., Carlson, M.: Physically based deformable models in computer graphics. Computer Graphics Forum 25, 809–836 (2006)
2. Paloc, C., Bello, F., Kitney, R.I., Darzi, A.: Online multiresolution volumetric mass spring model for real time soft tissue deformation. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002, Part II. LNCS, vol. 2489, pp. 219–226. Springer, Heidelberg (2002)

3. Lasseter, J.: Principles of traditional animation applied to 3d computer animation. ACM Siggraph Computer Graphics 21(4), 35–44 (1987)
4. Adams, B., Ovsjanikov, M., Wand, M., Seidel, H.P., Guibas, L.J.: Meshless modeling of deformable shapes and their motion. In: ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 77–86 (2008)
5. Diziol, R., Bender, J., Bayer, D.: Volume conserving simulation of deformable bodies. In: Eurographics 2009-Short Papers, pp. 37–40 (2009)
6. Müller, M., Heidelberger, B., Hennix, M., Ratcliff, J.: Position based dynamics. Journal of Visual Communication and Image Representation 18(2), 109–118 (2007)
7. Provot, X.: Deformation constraints in a mass-spring model to describe rigid cloth behaviour. In: Graphics Interface, pp. 147–155 (1995)
8. Mollemans, W., Schutyser, F., Van Cleynenbreugel, J., Suetens, P.: Fast soft tissue deformation with tetrahedral mass spring model for maxillofacial surgery planning systems. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3217, pp. 371–379. Springer, Heidelberg (2004)
9. Barauskas, R., Gulbinas, A., Barauskas, G.: Investigation of radiofrequency ablation process in liver tissue by finite element modeling and experiment. Medicina 43(4), 310–325 (2007)
10. Li, W., Luo, X., Johnson, A., Hill, N., Bird, N., Chin, S.: One-dimensional models of the human biliary system. Journal of Biomechanical Engineering 129(2), 164–173 (2007)
11. Lloyd, B., Székely, G., Harders, M.: Identification of spring parameters for deformable object simulation. IEEE Transactions on Visualization and Computer Graphics 13(5), 1081–1094 (2007)
12. Peterlík, I., Duriez, C., Cotin, S.: Modeling and real-time simulation of a vascularized liver tissue. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 50–57. Springer, Heidelberg (2012)
13. Irving, G., Schroeder, C., Fedkiw, R.: Volume conserving finite element simulations of deformable models. ACM Transactions on Graphics (TOG) 26(3), 13 (2007)

Adaptive Confidence Regions of Motion Predictions from Population Exemplar Models

Golnoosh Samei[1], Grzegorz Chlebus[2], Gabor Székely[1], and Christine Tanner[1]

[1] Computer Vision Laboratory, ETH Zurich,
Sternwartstrasse 7, 8092 Zurich, Switzerland
{samei,tannerch,szekely}@vision.ee.ethz.ch
[2] Faculty of Electronics, Janiszewskiego 11/17
Wroclaw University of Technology, 50-372 Wroclaw, Poland
gchlebus@gmail.com

Abstract. Precise radiation therapies require not only accurate prediction of the motion of the structures in the treatment region, but also confidence values of these predictions to enable planning of residual motion and detection of failure predictions. While various motion models have been proposed for the prediction of motion in the abdomen due to free-breathing, none has provided confidence regions. In this study we use the conditional probability density function of statistical liver motion models for predicting confidence regions, propose a method for optimizing the accuracy of the confidence regions and show the adaptability of the confidence regions due to partial observations when using exemplar models. The average accuracy of the confidence regions of single Gaussian (SG) models could be improved to the level of the exemplar models. Exemplar models provided on average better motion predictions (1.14 mm) and slightly smaller 68% confidence regions (1.36 mm) than the SG models (1.21 mm, 1.43 mm resp.). The confidence region size correlated temporally on average weakly (r=0.35) with the errors of the motion prediction for the exemplar models, leading to a higher percentage of treatable locations and lower motion prediction errors per duty cycle than SG models.

Keywords: Statistical population model, motion prediction, confidence regions, exemplar models.

1 Introduction

Minimal invasive tumour therapies are getting ever more sophisticated with novel treatment approaches and new devices allowing for improved targeting precision. Applying these effectively requires precise localization of the structures of interest. Respiration induces organ motion in the abdomen which cannot be neglected during therapy. Motion models have been developed for predicting the organ position from partial observations and mean prediction accuracies in the range of 1.1 to 4.6 mm have been reported for lungs and livers [1,2]. Yet no attempt has so far been made to estimate the accuracy of the motion prediction at the same

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 231–240, 2013.

time. The benefit of such estimates, for example in the form of local confidence regions, is twofold. First, prior to therapy, they allow for a more accurate planning by including the remaining motion uncertainty into the calculation of safety margins [3]. Second, the size of the confidence region could be used to pause the treatment when it is above a designated threshold to avoid high errors, which may happen due to unusual respiration patterns like coughing. Only temporally varying confidence regions can provide such an adaptive behaviour.

Analysis of the error statistics from the validation of the motion model could provide local confidence regions similar to methods used for shape estimations [4,5]. Yet, these are not adaptive to the current therapy situation. For statistical models, confidence regions can be derived from the model itself [6,7], instead of employing the error statistics. For example, the covariance matrix of a single Gaussian (SG) distribution model reflects the variation of the training data and hence the remaining uncertainty when using the mean of this distribution as the most likely prediction. Using such a model, confidence regions were estimated around the shapes predicted from partial observations [7]. Yet, for a single distribution the covariance remains the same for all predictions.

An exemplar model, consisting of a mixture of subject-specific Gaussian models, was proposed for predicting the liver motion during free-breathing [8]. It showed improved prediction performance over a SG model and has the advantage that the covariance matrix of the mixture distribution adapts to partial observations as the contribution of the individual Gaussian models is changing over time.

The contribution of this study is three-fold. First, we estimate confidence regions of motion predictions for the liver during free-breathing from statistical population models based on a SG or an exemplar model. Second, we propose a method for improving the accuracy of the confidence regions. Third, we show the advantage of the adaptability of the confidence regions for the exemplar model.

2 Material

Free-breathing 4D-MRI [9] data were obtained from 12 healthy volunteers (6 female, 6 male, mean age 31) over 40-75min. It consist of 25-30 slices with a spatial resolution of $1.8\times1.8\times4\text{mm}^3$ and a temporal resolution of 290-410 ms.

An intensity-based non-rigid registration was used [9] to determine the motion and deformation of the liver captured in the 4D-MRIs. A number of anatomically and biomechanically corresponding points were chosen manually for each subject to establish inter-subject correspondences. The positions of 290 corresponding points in the liver were obtained by performing cubic interpolation between these landmarks [9].

The position of 3 of these points, which were assumed to be feasible to track, were used as surrogates. These 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.

Using the same notation as in [8], the liver position was described by N points, where $p_t^i = [p_{SI_t}^i, p_{AP_t}^i, p_{LR_t}^i]$ denotes the position of the ith point at

time step t in the superior-inferior (SI), anterior-posterior (AP), and left-right (LR) direction. The whole liver position is then described by the $3N$ dimensional vector $\mathbf{p}_t$ formed from concatenating the $\mathbf{p}_t^i$ vectors. The motion of vector $\mathbf{p}_t$ is defined as $\Delta\mathbf{p}_t = \mathbf{p}_t - \mathbf{p}_{ref}$ with $\mathbf{p}_{ref}$ being the reference exhale vector.

3 Method

3.1 Motion Model

Subject-Specific SG Model. Assuming that $\Delta\mathbf{p}_t, t = 1..T$, with T being the number of time steps, belong to a $3N$ dimensional Gaussian distribution $\Delta\mathbf{p}_t \sim \mathcal{N}(\mu, \Sigma)$, the task is to find the most probable vector $\Delta\hat{\mathbf{p}}_t$, given a subset of its elements $\mathbf{s}_t$, called surrogate. Decomposing $\Delta\mathbf{p}_t$, mean μ, and covariance matrix Σ into the components relating to surrogate $\mathbf{s}_t$ and to the rest of the points we wish to predict ($\mathbf{r}_t$), we get $\Delta\mathbf{p}^T = \left[\mathbf{s}^T, \mathbf{r}^T\right]$, $\mu^T = \left[\mu_\mathbf{s}^T, \mu_\mathbf{r}^T\right]$, and $\Sigma = \begin{bmatrix} \Sigma_{\mathbf{ss}} & \Sigma_{\mathbf{sr}} \\ \Sigma_{\mathbf{rs}} & \Sigma_{\mathbf{rr}} \end{bmatrix}$. If the distribution of $\Delta\mathbf{p}_t$ is Gaussian, then the conditional distribution $(\Delta\mathbf{p}_t|\mathbf{s}_t)$ is also a Gaussian distribution of the form [10],

$$(\Delta\mathbf{p}_t|\mathbf{s}_t) \sim \mathcal{N}\left(\mu + \begin{bmatrix} \Sigma_{\mathbf{ss}} \\ \Sigma_{\mathbf{rs}} \end{bmatrix} {\Sigma_{\mathbf{ss}}}^{-1}(\mathbf{s}_t - \mu_\mathbf{s}), \Sigma - \begin{bmatrix} \Sigma_{\mathbf{ss}} \\ \Sigma_{\mathbf{rs}} \end{bmatrix} {\Sigma_{\mathbf{ss}}}^{-1} \begin{bmatrix} \Sigma_{\mathbf{ss}} \\ \Sigma_{\mathbf{rs}} \end{bmatrix}^T\right). \quad (1)$$

Therefore, the most probable vector $\Delta\hat{\mathbf{p}}_t$ given $\mathbf{s}_t$, is the mean of $\Delta\mathbf{p}_t|\mathbf{s}_t$:

$$\Delta\hat{\mathbf{p}}_t = \mu + \begin{bmatrix} \Sigma_{\mathbf{ss}} \\ \Sigma_{\mathbf{rs}} \end{bmatrix} {\Sigma_{\mathbf{ss}}}^{-1}(\mathbf{s}_t - \mu_\mathbf{s}). \quad (2)$$

Population SG Model. Given a dataset with J subjects, where each subject $j = 1...J$, has T_j observations of the same N points, a population SG model is built 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.

Population Exemplar Model. To create an exemplar model [8], a subject-specific SG model M^j was built for each subject j. To predict $\Delta\mathbf{p}_t$, for a new subject, given $\mathbf{s}_t$, the motion vector predictions $\Delta\hat{\mathbf{p}}_t^j$ are obtained for all M^j models by Eq. (2). These predictions are then combined by a distance-weighted k-nearest-neighbor approach [11], using the squared Mahalanobis distance between $\mathbf{s}_t$ and the corresponding mean surrogate observation of M^j:

$$d(\mathbf{s}_t, M^j) = (\mathbf{s}_t - \mu_\mathbf{s}^j)^T {\Sigma_{\mathbf{ss}}^j}^{-1}(\mathbf{s}_t - \mu_\mathbf{s}^j).$$

The prediction from model M^j at time t is then weighted by w_t^j, which is computed from the normalized inverse of the above distance:

$$w_t^j = \frac{1/(d(\mathbf{s}_t, M^j) + \eta)}{\sum_{k=1}^{J} 1/(d(\mathbf{s}_t, M^k) + \eta)}, \quad (3)$$

where η is a small positive value to avoid division by 0. Finally $\Delta\hat{\mathbf{p}}_t$ is given by

$$\Delta\hat{\mathbf{p}}_t = \sum_{j=1}^{J} w_t^j \Delta\hat{\mathbf{p}}_t^j. \tag{4}$$

3.2 Estimation of Confidence Regions

Besides predicting the current liver motion $\Delta\hat{\mathbf{p}}_t$, we would also like to estimate the confidence region $C_\alpha(\Delta\hat{\mathbf{p}}_t^i)$ around $\Delta\hat{\mathbf{p}}_t^i$, with a probability of α (i.e. confidence level) to include the true motion vector $\Delta\mathbf{p}_t^i$.

The prediction of $\Delta\hat{\mathbf{p}}_t$ for the SG population model is based on the mean of the conditional distribution $(\Delta\mathbf{p}_t|\mathbf{s}_t)$ defined in Eq. (1). The covariance matrix of this distribution:

$$\Sigma_{\Delta\mathbf{p}|\mathbf{s}} = \left[\Sigma - \begin{bmatrix}\Sigma_{\mathbf{ss}}\\ \Sigma_{\mathbf{rs}}\end{bmatrix} {\Sigma_{\mathbf{ss}}}^{-1} \begin{bmatrix}\Sigma_{\mathbf{ss}}\\ \Sigma_{\mathbf{rs}}\end{bmatrix}^T\right], \tag{5}$$

provides an estimate of the model uncertainty which is independent of t.

For the exemplar model, the motion prediction $\Delta\hat{\mathbf{p}}_t$ is based on the weighted linear combination of the motion predictions from the subject-specific SG models $(\Delta\hat{\mathbf{p}}_t^j)$, see Eq. (4). Since the mth moment about zero of a mixture model is simply a weighted average of the mth moments of the mixture components, we can write the mean and covariance of the corresponding mixture conditional distribution around $\Delta\mathbf{p}_t$, as follows by [6]

$$\mu_{\Delta\mathbf{p}_t|\mathbf{s}_t} = \sum_{j=1}^{J} w_t^j \mu_{\Delta\mathbf{p}_t^j|\mathbf{s}_t^j} = \sum_{j=1}^{J} w_t^j \Delta\hat{\mathbf{p}}_t^j = \Delta\hat{\mathbf{p}}_t \tag{6}$$

$$\begin{aligned}\Sigma_{\Delta\mathbf{p}_t|\mathbf{s}_t} &= \sum_{j=1}^{J} w_t^j \left(\Sigma_{\Delta\mathbf{p}^j|\mathbf{s}^j} + \mu_{\Delta\mathbf{p}_t^j|\mathbf{s}_t^j}\mu_{\Delta\mathbf{p}_t^j|\mathbf{s}_t^j}{}^T\right) - \mu_{\Delta\mathbf{p}_t|\mathbf{s}_t}\mu_{\Delta\mathbf{p}_t|\mathbf{s}_t}^T \\ &= \sum_{j=1}^{J} w_t^j \left(\Sigma_{\Delta\mathbf{p}^j|\mathbf{s}^j} + \Delta\hat{\mathbf{p}}_t^j \Delta\hat{\mathbf{p}}_t^{j\,T}\right) - \Delta\hat{\mathbf{p}}_t \Delta\hat{\mathbf{p}}_t{}^T.\end{aligned} \tag{7}$$

Note that $\Sigma_{\Delta\mathbf{p}_t|\mathbf{s}_t}$ is dependent on t. Point-wise 3x3 covariance matrices $\Sigma_{\Delta\mathbf{p}_t^i|\mathbf{s}_t}$ can be extracted from $\Sigma_{\Delta\mathbf{p}_t|\mathbf{s}_t}$ for estimating point-wise confidence regions $C_\alpha(\hat{\mathbf{p}}_t^i)$.

Confidence region $C_\alpha(\hat{\mathbf{p}}_t^i)$ is the interior of a 3D ellipsoid centered at $\hat{\mathbf{p}}_t^i$ with axes defined by the eigenvalue decomposition of $\Sigma_{\Delta\mathbf{p}_t^i|\mathbf{s}_t}$ and a size such that the probability $P(\mathbf{p} \in C_\alpha(\hat{\mathbf{p}}_t^i)) = \alpha$. Using the approach of Blanc et al. [5], the boundary of this ellipsoid is defined by the set of 3D points ρ which have a constant Mahalanobis distance D_α to the predicted points $\hat{\mathbf{p}}_t^i$:

$$D_\alpha = \sqrt{(\rho - \hat{\mathbf{p}}_t^i)^T \Sigma^{-1}_{\Delta\mathbf{p}_t^i|\mathbf{s}_t} (\rho - \hat{\mathbf{p}}_t^i)}, \tag{8}$$

where $D_\alpha^2 = K^{-1}(\alpha, 3)$ and $K^{-1}(\alpha, 3)$ stands for the inverse of the cumulative χ^2 distribution function with 3 degrees of freedom.

Common margin recipes employed in radiotherapy assume no correlations between the error components in the different spatial directions [3]. To fulfill this assumption, we calculate confidence regions based on only the diagonal of $\Sigma_{\Delta\mathbf{p}_t^i|\mathbf{s}_t}$, i.e. assuming zero covariance.

We calculated 68% confidence regions for usage as random motion error (σ_m) in the margin recipe of van Herk et al. [12,3],

$$\nu = 1.64(\sqrt{(\sigma_m^2 + \sigma_s^2 + \sigma_p^2)} - \sigma_p^2), \tag{9}$$

where σ_s is the set-up error, σ_p is the width of the beam penumbra and ν is the recommended safety margin to achieve a minimum dose to the clinical target volume of 95% of the nominal dose. Furthermore, we assessed also the performance for all confidence levels by determining the area between nominal and effective confidence curves.

3.3 Optimization of Confidence Regions

The theoretic confidence regions calculated in Sec. 3.2 are only estimates and might not provide the nominal confidence levels. Therefore it is important to estimate the effective confidence level [5]. For the motion predicted for each point i, subject j and time t, $\Delta\hat{\mathbf{p}}_t^{i,j}$, the α-confidence region $C_\alpha(\Delta\hat{\mathbf{p}}_t^{i,j})$ can be computed as described above. The probability that the true motion $\Delta\mathbf{p}^{i,j}$ is within this region can be estimated by the effective frequency $\hat{\alpha}^{i,j}$ with which this happens for the T_j time frames of subject j:

$$\hat{\alpha}^{i,j}(\alpha) = \frac{1}{T_j}\sum_{t=1}^{T_j} \mathbf{1}_{C_\alpha(\Delta\hat{\mathbf{p}}_t^{i,j})}(\Delta\mathbf{p}_t^{i,j}) \lim_{T_j\to\infty} P(\Delta\mathbf{p}^{i,j} \in C_\alpha(\Delta\hat{\mathbf{p}}^{i,j})), \tag{10}$$

where $\mathbf{1}_A(x)$ is the indicative function, which is 1 if $x \in A$ and otherwise 0. For all points of subject j, the effective value $\hat{\alpha}^j(\alpha) = 1/N\sum_{i=1}^N \hat{\alpha}^{i,j}$. The accuracy of the estimated confidence regions can then be evaluated by comparing the effective values $\hat{\alpha}^{i,j}(\alpha)$ and $\hat{\alpha}^j(\alpha)$ to the corresponding nominal value α.

Furthermore, as we have motion observations of a whole population, we can try to improve the accuracy of the confidence region estimations by learning the relationship f between the nominal and the effective confidence level using subject-wise optimization ($f_j(\alpha)$), and point- and subject-wise optimization $f_{i,j}(\alpha)$) in advance from the population. Assuming equal cost for over- and under-estimation of the confidence level and using nested cross-validation, the effective confidence level of the left-out case u is approximated to be the mean effective confidence level of the population V (e.g. $f_u(\alpha) = \sum_{v\in V}\hat{\alpha^v}(\alpha)/(J-1)$), see Algorithm 1.

4 Results

We trained population SG and exemplar models to predict the motion and the associated confidence regions of the liver using cross-validation, see Algorithm 1.

Data: motion data for subject set U={1,2,...,J}
foreach $u \in U$ **do**
  $V = U/u$;
  foreach $v \in V$ **do**
    $W = V/v$;
    train population model on W;
    foreach $\alpha \in [0, 1]$ **do**
      for $i = 1 \rightarrow N$ **do**
        for $t = 1 \rightarrow T_v$ **do**
          predict $\Delta\hat{\mathbf{p}}_t^{i,v}$ and confidence region $C_\alpha(\Delta\hat{\mathbf{p}}_t^{i,v})$;
          determine $\mathbf{1}_{C_\alpha(\Delta\hat{\mathbf{p}}_t^{i,v})}(\Delta\mathbf{p}_t^{i,v})$;
        end
        calculate effective confidence level $\hat{\alpha}^{i,v}(\alpha)$;
      end
      calculate effective confidence level $\hat{\alpha}^{v}(\alpha)$;
    end
  end
  train population model on V;
  foreach $\alpha \in [0, 1]$ **do**
    estimate $f_u(\alpha) = \sum_{v \in V} \hat{\alpha}^{v}(\alpha)/(J-1)$;
    for $i = 1 \rightarrow N$ **do**
      estimate $f_{i,u}(\alpha) = \sum_{v \in V} \hat{\alpha}^{i,v}(\alpha)/(J-1)$;
      for $t = 1 \rightarrow T_u$ **do**
        predict $\Delta\hat{\mathbf{p}}_t^{i,u}$ and confidence region $C_\alpha(\Delta\hat{\mathbf{p}}_t^{i,u})$;
        determine $\mathbf{1}_{C_\alpha(\Delta\hat{\mathbf{p}}_t^{i,u})}(\Delta\mathbf{p}_t^{i,u})$;
      end
      calculate effective confidence level $\hat{\alpha}^{i,u}(\alpha)$;
    end
  end
  foreach $\alpha \in [0, 1]$ **do**
    for $i = 1 \rightarrow N$ **do**
      error no optimization: $E_0^{i,u}(\alpha) = \hat{\alpha}^{i,u}(\alpha) - \alpha$;
      error subject-wise optimization: $E_{LS}^{i,u}(\alpha) = \hat{\alpha}^{i,u}(f_u^{-1}(\alpha)) - \alpha$;
      error point- & subject-wise opt.: $E_{LP}^{i,u}(\alpha) = \hat{\alpha}^{i,u}(f_{i,u}^{-1}(\alpha)) - \alpha$;
    end
  end
end

Algorithm 1: Method for learning and evaluating the relationship between nominal and effective confidence level ($f_j(\alpha)$, $f_{i,j}(\alpha)$)

Accuracy of Confidence Regions. Fig. 1 shows the relationship between the mean nominal and effective confidence level per subject and for all subjects. General underestimation and a larger variation can be observed for the SG model, while the exemplar model is relatively balanced.

Learning the relationship between the nominal and effective confidence level improved the median error at the 68% confidence level (from -0.22 to 0.03) and overall (median area between curves from 0.17 to 0.12) for the SG model, see Fig. 2 (a,c). No such improvement can be observed for the exemplar model, which has a similar performance even unoptimized.

Motion Prediction and Confidence Regions. The SG model had higher motion prediction errors and required larger 68% and 95% mean confidence regions than the exemplar model, see Table 1. The size of the confidence region

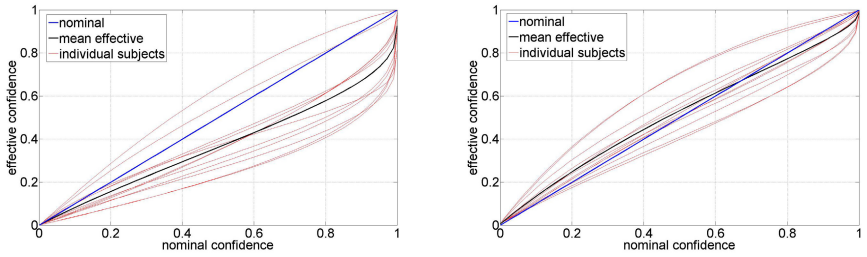


Fig. 1. Mean nominal versus effective confidence level for all points per subject and over all subjects for (left) the SG model and (right) the exemplar model

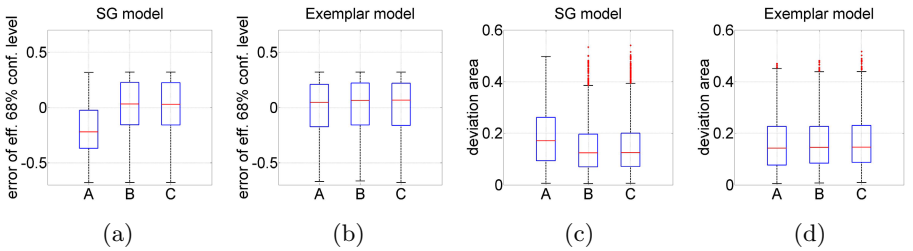


Fig. 2. Boxplots illustrating the distribution of (a,b) error of effective 68% confidence level ($E^{i,j}(0.68)$), (c,d) area between the nominal and effective confidence curves ($A^{i,j} = \int_{\alpha\in[0,1]} |E^{i,j}(\alpha)|$) for all points and subjects for (a,c) SG model and (b,d) exemplar model (A) without optimization (E_0), (B) subject-wise optimized (E_{LS}), and (C) point- and subject-wise optimization (E_{LP})

correlated temporally weakly with the absolute motion prediction error. No such correlation exists by definition for the static confidence regions of the SG model.

Fig. 3 illustrates for a single point the motion prediction with estimated 68% confidence regions over time. It can be observed that the size of the confidence regions remains the same over time for the SG model, while the exemplar model has smaller confidence regions at end exhalation (motion around 0 mm) than at end inhalation. The variation in confidence region size and shape for neighbouring points is illustrated in Fig. 4. For this subject, the SG model results are relatively similar to the mean confidence regions of the exemplar model over time (Fig. 4a,b). Yet adaption of the exemplar model to particular states, like end inhalation and exhalation, results in quite different confidence regions (Fig. 4c,d). On average, confidence regions are larger in the anterior-inferior (left-bottom) region where drift (change in end exhalation position) is occurring [9].

Clinical Significance of the Exemplar Models. To evaluate the advantages of the exemplar method, we assumed the following clinical scenario. Given the clinicians only accept margins of at most threshold θ, i.e. treatment is applied only if $\nu_t^{i,j} < \theta$ in all directions (see Eq. 9).

Table 1. Summary of distribution of motion prediction error (Euclidean distance) and size of optimized 68% and 95% confidence regions (in mm) for all points and subjects after averaging over all time frames. Mean temporal correlation (r) between size of confidence region and absolute motion prediction error in each direction.

	SG model				Exemplar model				
	Mean	Median	SD	95%	Mean	Median	SD	95%	r
Motion prediction error	1.21	1.00	0.86	2.90	1.14	0.93	0.82	2.77	n/a
68% Confidence region AP	1.16	1.15	0.46	1.94	1.10	0.94	0.67	2.39	0.38
68% Confidence region SI	1.56	1.51	0.60	2.66	1.50	1.30	0.89	3.22	0.36
68% Confidence region LR	1.56	1.51	0.59	2.61	1.48	1.31	0.83	3.08	0.31
95% Confidence region AP	2.26	2.24	0.89	3.77	1.91	1.63	1.16	4.17	0.36
95% Confidence region SI	3.03	2.93	1.17	5.16	2.62	2.26	1.56	5.61	0.35
95% Confidence region LR	3.03	2.94	1.15	5.07	2.59	2.27	1.44	5.38	0.30

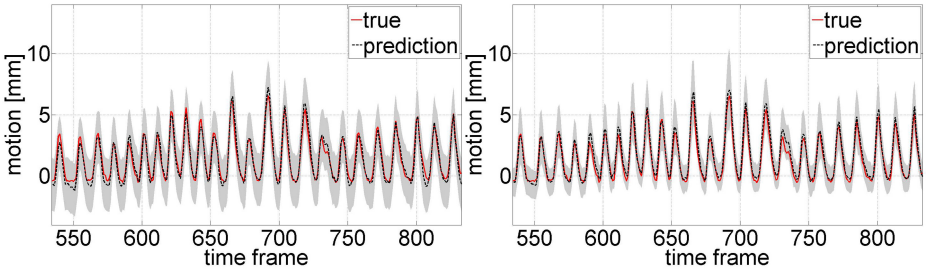


Fig. 3. Plots of 300 time frames from subject 2 and point 58 (marked by 'x' in Fig. 4), illustrating true motion, predicted motion and 68% confidence region (gray region) after learning the effective confidence level for the SI direction for (left) SG model and (right) exemplar model

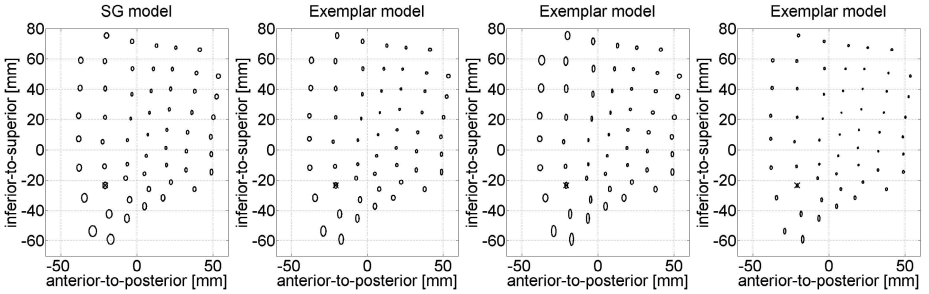


Fig. 4. Illustration of the 68% confidence regions of the points on a central sagittal liver slice for patient 2 for (a) SG model, (b-d) exemplar model (b) on average, (c) at end inhalation ($t = 800$) and (d) at end exhalation ($t = 700$). See Fig. 3 for confidence regions over time of the point marked by 'x'.

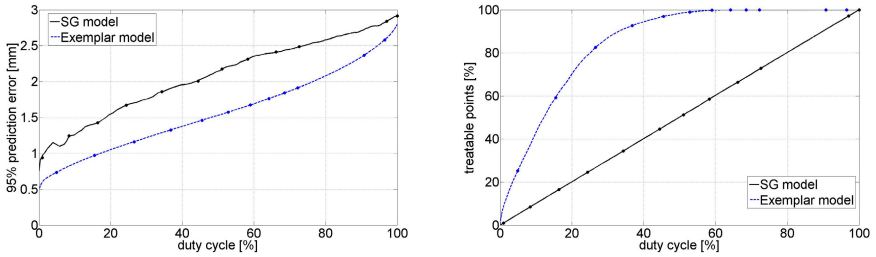


Fig. 5. Duty cycle versus (left) 95 percentile prediction error and (right) percentage of treatable points over all points and subjects for SG (solid black line) and exemplar model (dotted blue line). Thresholds [0.1, 0.2, ..., 1.0, 2.0, 3.0]mm are marked by dots.

Then we calculated for each method the duty cycle (percentage of treatment to overall time), the motion prediction error during treatment, and the percentage of points which had some treatment (called treatable points). Note that for the fixed confidence region size of the SG model, it will not be possible to treat certain points at any given time, while others will always be included. Fig.5 (left) shows that for all duty cycles the exemplar model achieves a lower 95% prediction error. The right plot shows that for each duty cycle much more locations are treatable using the exemplar models versus the SG model.

5 Conclusions

We proposed a method for providing optimized confidence regions of the motion predictions from statistical models. It is based on the covariance matrix of the model's conditional distribution and on learning the relationship between the nominal and effective significance level. For the exemplar model, this covariance matrix is changing due to the varying contributions from the subject-specific SG models for different partial observations. Such adaptability during therapy is a pre-requisite for accurate margin calculations and for detecting motion not covered by the model, such that treatment can be paused to avoid large errors.

Optimization of the confidence regions was important for the SG model, while the exemplar model provided already a balanced relationship. Motion prediction errors and confidence regions were on average smaller for the exemplar model than the SG model. Exemplar models provided lower 95% prediction errors and more treatable locations per duty cycle. Further improvements may come from using the whole covariance matrices, having a larger population, and splitting respiratory and drift motion. The probabilistic approach allows for integration of similar confidence measures from other components of the system (e.g. surrogate tracking, temporal prediction) to ultimately enable a safe and accurate therapy.

Acknowledgments. We would like to acknowledge the funding from the Zeno Karl Schindler Foundation and 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 (2012)
2. McClelland, J., Hawkes, D., Schaeffter, T., King, A.: Respiratory motion models: A review. Medical Image Analysis 17, 19–42 (2012)
3. Van Herk, M.: Errors and margins in radiotherapy. In: Seminars in Radiation Oncology, vol. 14, pp. 52–64. Elsevier (2004)
4. Blake, A., Isard, M.: Active contours (1998)
5. Blanc, R., Syrkina, E., Székely, G.: Estimating the confidence of statistical model based shape prediction. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) IPMI 2009. LNCS, vol. 5636, pp. 602–613. Springer, Heidelberg (2009)
6. Bishop, C.M.: Pattern recognition and machine learning, vol. 1. Springer, New York (2006)
7. Baka, N., de Bruijne, M., Reiber, J., Niessen, W., Lelieveldt, B.: Confidence of model based shape reconstruction from sparse data. In: Int. Symposium on Biomedical Imaging: From Nano to Macro, pp. 1077–1080. IEEE (2010)
8. Samei, G., Tanner, C., Székely, G.: Predicting liver motion using exemplar models. In: Yoshida, H., Hawkes, D., Vannier, M.W. (eds.) Abdominal Imaging 2012. LNCS, vol. 7601, pp. 147–157. Springer, Heidelberg (2012)
9. von Siebenthal, M., Székely, G., Lomax, A., Cattin, P.C.: 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)
10. Ahrendt, P.: The multivariate gaussian probability distribution. Tech. rep. (2005)
11. Dudani, S.A.: The distance-weighted k-nearest-neighbor rule. Transactions on Systems, Man and Cybernetics (4), 325–327 (1976)
12. Van Herk, M., Remeijer, P., Rasch, C., Lebesque, J.V.: The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. Int. J. Radiat. Oncol. Boil. Phys. 47(4), 1121 (2000)

A Generic, Robust and Fully-Automatic Workflow for 3D CT Liver Segmentation

Romane Gauriau[1,2], Rémi Cuingnet[2], Raphael Prevost[2], Benoit Mory[2], Roberto Ardon[2], David Lesage[2], and Isabelle Bloch[1]

[1] Institut Mines-Telecom, Telecom ParisTech, CNRS LTCI, 46 rue Barrault, 75013 Paris, France
[2] Philips Research MediSys, 33 rue de Verdun, 92156 Suresnes Cedex, France

Abstract. Liver segmentation in 3D CT images is a fundamental step for surgery planning and follow-up. Robustness, automation and speed are required to fulfill this task efficiently. We propose a fully-automatic workflow for liver segmentation built on state-of-the-art algorithmic components to meet these requirements. The liver is first localized using regression forests. A liver probability map is computed, followed by a global-to-local segmentation strategy using a template deformation framework. We evaluate our method on the SLIVER07 reference database and confirm its state-of-the-art results on a large, varied database of 268 CT volumes. This extensive validation demonstrates the robustness of our approach to variable fields of view, liver contrast, shape and pathologies. Our framework is an attractive tradeoff between robustness, accuracy (mean distance to ground truth of 1.7mm) and computational speed (46s). We also emphasize the genericity and relative simplicity of our framework, which requires very limited liver-specific tuning.

Keywords: Liver segmentation, fully-automatic segmentation, template deformation, regression forest, 3D-CT.

1 Introduction

Liver segmentation is required in many clinical contexts such as tumor resection, follow-up or liver transplantation. It enables the computation of anatomical measures that are important for clinical diagnosis, surgery planning and radiation dose calculation [1]. Manual liver segmentation in 3D is both tedious and time-consuming and its automation is particularly challenging given the high variability of liver shapes, pathologies and contrast in different CT phases.

The literature on liver segmentation includes a large variety of interactive, semi-automatic and automatic methods. Due to space restrictions, we refer the reader to recent and extensive reviews [2,3] and to the SLIVER07 segmentation challenge [4,5] for more detailed bibliographic overviews. Those reviews highlight different groups of methods such as intensity-based, active contours, statistical shape models, graph-cuts and atlas-based registration. They also show that most segmentation methods focus on CT images with specific fields of view and particular CT phases (often contrast-enhanced) which may restrict their clinical use.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 241–250, 2013.

Quantitative evaluation is another key point when comparing different methods. The challenge SLIVER07 [4] has become a reference for liver segmentation, enabling fast and easy comparisons [5]. Unfortunately this database is limited to 20 training and 10 testing datasets. In the literature, only few methods were validated on extensive databases. One can cite [6,7,8], where the authors used proprietary databases composed of 277, 75 and 48 images respectively. However the differences in evaluation criteria and database composition make comparisons of these methods difficult.

We propose a fully-automatic and robust method for the segmentation of the liver on CT data. Our workflow is inspired by [9], which proved its efficiency for CT kidney segmentation. Our method consists of four main parts built on state-of-the-art algorithmic components as shown in Fig. 1. The liver and the heart are first localized using regression forests [10] (see Section 2). We then compute a liver probability map based on intensity distributions (see Section 3). A template deformation framework [11] performs the liver segmentation using a global-to-local strategy (see Section 4.2). A final refinement step is applied using the original image.

In Section 5, we reuse and extend the validation framework of the SLIVER07 challenge and present a quantitative evaluation on both the SLIVER07 database and a large and diverse database (268 CT volumes with various fields of view, contrasts, liver shapes and pathologies). The results demonstrate, in an extensive and coherent fashion, the computational efficiency, robustness and accuracy of our method.

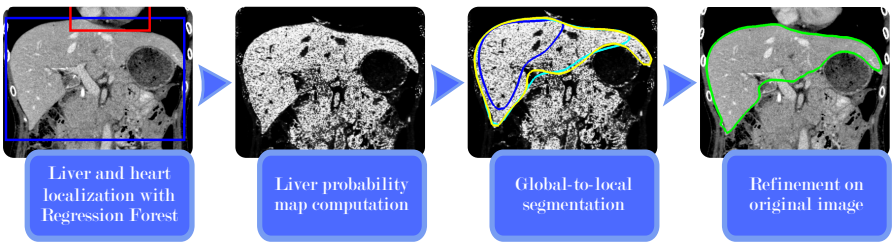


Fig. 1. Workflow of our fully-automatic liver segmentation method

2 Liver Localization Using Regression Forest

The authors of [10] recently demonstrated that regression forests are robust and memory efficient for the quick localization of multiple abdominal organs in 3D CT scans (1s/tree with their C++ implementation). They compared this method with the commonly used atlas-based method [12] and demonstrated that it is about a hundred times faster, uses ten times less memory and is more accurate.

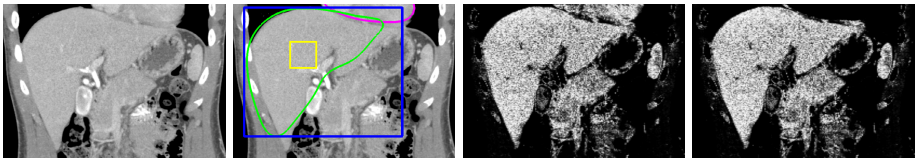
The main idea behind regression forest is to use random forest for non-linear regression of multidimensional output given multidimensional input. Using training data, binary trees are built so as to split the data into clusters, in which the prediction can be achieved with a simple regression function.

As in [9], we use the regression forest method to detect bounding boxes of the liver and heart (see Fig. 1), each of them being parametrized with a vector of $\mathbb{R}^6$ composed of the coordinates of two extremal vertices. The training phase is performed using random subsets of voxels of the training images. The features used are the same as in [10]: for each voxel we compute the mean intensities in two randomly displaced boxes. This exploits the fact that the intensities in CT images have a real physical meaning. In the testing phase a random selection of voxels votes for the predicted labels. A detailed and comprehensive description of the regression forests method can be found in [10].

This approach provides robust estimates of the positions and sizes of the liver and heart which are used to derive a liver probability map described hereafter.

3 Liver Probability Map Computation

Segmenting the liver directly in the image may provide insufficient results, in particular in images with poor contrast and fuzzy liver contours. Consequently, we propose to also take advantage of intensity distribution to pre-process the image and enhance liver voxels as shown in Fig. 2.



(a) (b) (c) (d)

Fig. 2. Probability map computation steps: (a) original image, (b) liver predicted box (blue), fitted mean liver (green), heart segmentation (pink), voxels patch (yellow), probability map (c) before and (d) after heart masking

3.1 Fitting a Mean Liver Model in the Predicted Bounding Box

Mean Liver Computation. A mean liver model is built using a set of manually segmented liver shapes, represented as meshes. We first register the shape meshes using the fast and robust registration method of [13]. The main interest of this method over the classical Iterative Closest Point approach [14] is to overcome the problem of point correspondence and to use a robust norm to obtain a consistent registration of the shapes. The mean liver model is obtained by averaging the implicit functions of the registered shapes.

Mean Liver Fitting. The mean liver shape is scaled anisotropically so as to best fit the predicted bounding box of the liver.

3.2 Estimation of the Intensity Histogram

From the previously fitted mean liver model barycenter we select a cuboidal patch ($\sim 60^3$ mm^3) of voxels, as illustrated in Fig. 2(b). The intensity histogram is computed in this patch. Its normalization provides a function $h : \mathbb{R} \rightarrow [0, 1]$ which gives an estimation of the probability of each intensity value to belong to the liver.

3.3 Coarse Segmentation of the Heart

In addition to the estimation of liver intensities, we roughly segment the heart in the image. This segmentation is performed on the original image with the template deformation framework described in Subsection 4.1. We initialize the algorithm with a mean heart model, built similarly to the mean liver shape and fitted in the predicted heart bounding box (see Subsection 3.1). Deformation parameters (see Table 1) are set so as to prevent the heart contour from leaking in the liver. Conversely the rough binary mask M_h we obtain will prevent the liver contour from leaking in the heart in the subsequent steps.

3.4 Probability Map Computation

The liver probability map $M_l : \Omega \rightarrow [0, 1]$ is defined as:

$$\forall \mathbf{x} \in \Omega, M_l(\mathbf{x}) = (1 - M_h(\mathbf{x}))\ h(I(\mathbf{x})) \tag{1}$$

where Ω is the image domain and I is the image. A subsampling of the image can be previously applied to increase the computational efficiency.

This probability map (see Fig. 2 (d)) is used in subsequent segmentation steps.

4 Template-Based Global-to-Local Segmentation

We use the template deformation framework introduced in [11] to extract the liver contours from the probability map and the original image. Hereafter we briefly describe the model-based deformation algorithm we employ and detail the proposed global-to-local strategy to complete the segmentation.

4.1 Template Deformation Framework

The model-based approach of [11] is specially suited when target objects have partially unclear edges as the algorithm fairly extrapolates the contours. Key advantages of the method are its speed, robustness and its ability to use both contour and region information. Let us present the main principles of this method.

Given an image $J : \Omega \rightarrow \mathbb{R}$ and an initial shape model represented by an implicit function $\phi : \Omega \rightarrow \mathbb{R}$, we look for the transformation $\psi : \Omega \rightarrow \Omega$ minimizing the following energy function:

$$E_s(\psi) = \alpha \underbrace{\int_{(\phi\circ\psi)^{-1}(0)} -\langle \vec{\nabla} J(\mathbf{x}), \vec{\mathbf{n}}(\mathbf{x})\rangle d\mathbf{x}}_{\text{flux term}} + (1-\alpha) \underbrace{\int_{(\phi\circ\psi)^{-1}(\mathbb{R}^+)} r(\mathbf{x}) d\mathbf{x}}_{\text{region term}} + \lambda \underbrace{\mathcal{R}(\psi)}_{\substack{\text{regularization} \\ \text{term}}} \tag{2}$$

where

- $\alpha \in [0,1]$ is a constant defining the relative influence of flux and region terms (it enables to define whether we want to rely more on image contours or more on intensity contrast between regions);
- $\vec{\nabla} J(\mathbf{x})$ is the gradient of the image J in $\mathbf{x}$ and $\langle\ .\ ,\ .\ \rangle$ is the scalar product;
- $\vec{\mathbf{n}}(\mathbf{x})$ is the normal vector to the shape at point $\mathbf{x}$;
- $r(\mathbf{x})$ is the region term defined as $r(\mathbf{x}) = log\frac{P_{\text{int}}(J(\mathbf{x}))}{P_{\text{ext}}(J(\mathbf{x}))}$ where P_{int} and P_{ext} are the intensities distributions inside and outside the deformed object regularly estimated on the working image;
- $\mathcal{R}(\psi)$ prevents large deviations from the original shape model;
- $\lambda \in [0,1]$ is a constant parameter tuning the strength of the shape constraint.

In the general formulation the transformation ψ is decomposed as $\psi = \mathcal{L} \circ \mathcal{G}$ where $\mathcal{G}$ is a global linear transformation and $\mathcal{L}$ is a non-rigid local transformation (refer to [11] for more details). Further we introduce $\mathcal{G}_r$ and $\mathcal{G}_s$ defining a rigid and a similarity transformation, respectively. The regularization term is defined as $\mathcal{R}(\psi) = \frac{1}{2}\|\mathcal{L} - Id\|_2^2$ where Id is the identity transformation.

4.2 Global-to-Local Segmentation Workflow

Liver segmentation is performed according to an original global-to-local strategy on the probability map using the above-mentioned algorithm. Four steps help refining the contour progressively: 3 steps are performed on the probability map M_l and the final step on the original image I as shown on Fig. 3. Parameters have been set experimentally and are kept identical in all processed examples (see Table 1). We observed a relatively low sensitivity to parameter variations in practice.

Table 1. Parameters used for global-to-local segmentation

	J	ψ	α	λ
Heart	I	$\mathcal{L} \circ Id$	0.8	0.1
Liver				
step 1	M_l	$Id \circ \mathcal{G}_r$	0.4	–
step 2	M_l	$\mathcal{L} \circ \mathcal{G}_s$	0.8	0.03
step 3	M_l	$\mathcal{L} \circ Id$	1	0.01

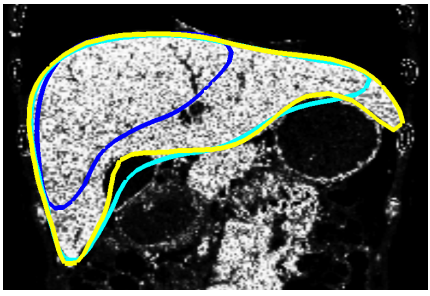


Fig. 3. The three segmentation steps on the probability map: step 1 (blue), step 2 (cyan), step 3 (yellow)

Step 1: Initialization. The liver shape model is fitted to the predicted liver bounding box. As the template deformation method tends to favor expansion displacements, we scale down the model by a factor 0.7. Then a first step aims at globally registering the shape model without any local deformation. Equation 2 is minimized using the parameters reported in Table 1. The high weight on the region term constrains the shape inside the liver, thus facilitating expansion in subsequent deformation steps.

Step 2: Coarse Segmentation. We deform the previous result, still minimizing Equation 2 but with different parameters (see Table 1) allowing some local deformations. The flux term is now more important than the region term so that the model contours match liver edges more accurately.

Step 3: Local Deformation. The third step helps refining the segmentation. Now we only optimize local deformations (see Table 1). The flux term is used alone in order to reach the contours. Releasing the shape constraint finally helps reaching stretched parts such as liver tips.

Step 4: Refinement on the Original Image. To improve the accuracy of the final segmentation, we apply guided filtering [15] on the binary mask of the previous segmentation result, using the original image as guide. As detailed in [15], the guided filter acts as a fast, local matting/feathering refinement step enabling the final segmentation to better match the edges of the original image.

5 Experiments and Results

In this part we present experiments on two databases. The first one is SLIVER07 [4], which has become a reference for liver segmentation evaluation. Secondly we use a large database of 268 diverse CT volumes, further demonstrating the accuracy and robustness of the method. In those two experiments both the regression forest and the mean liver model are learned solely on the 20 training datasets of SLIVER07 (publicly available), for the sake of result reproducibility. Computational times are given for a C++ implementation on a machine with four 2.3GHz cores (Core i7-2820QM) and 8Go RAM.

5.1 Training

For mean liver model building and regression forests training, we use the 20 training images of SLIVER07. Before regression forest training we do not pre-register or normalize the images. The forest of 7 trees and 12 decision levels is learned after randomly selecting a subset of 40.000 voxels per image. The minimum node size is 50 and the training computational time is about 10 minutes.

Table 2. The five best automatic methods on SLIVER07 database. We report the computational time (per image), the number of training volumes and the SLIVER07 measures. (n/a: non available)

	Time	Training volumes	OVE [%]	OVE Score	VOD [%]	VOD Score	AVD [mm]	AVD Score	RMSD [mm]	RMSD Score	MAXD [mm]	MAXD Score	Total Score
[16]	15min	122	**6.1**	**76.2**	−2.9	84.7	**0.95**	**76.3**	**1.9**	**74.0**	18.7	75.4	**77.3 ± 9.4**
[17]	3min	112	6.5	74.7	**1.0**	**86.4**	1.0	74.5	2.0	72.3	**18.3**	**75.9**	76.8 ± 3.8
[18]	n/a	n/a	6.4	75.1	2.3	85.0	1.0	74.9	1.9	73.4	20.8	72.7	76.2 ± 5.9
[19]	n/a	-	7.6	70.4	−1.3	85.4	1.3	68.0	2.4	67.4	22.1	70.9	72.4 ± 8.6
Ours	**46s**	**20**	7.2	71.7	2.6	85.0	1.3	67.0	2.6	64.2	23.1	69.6	71.5 ± 10.0

5.2 Evaluation on the SLIVER07 Database

We tested our method on the SLIVER07 challenge database [4] which is composed of 20 training and 10 testing 3D CT volumes (average slice and interslice resolution are 0.7mm±0.1 and 1.5mm±0.9, respectively) rather focused on the liver. The regression forest and mean liver shape were trained on the 20 training samples while we tested the algorithm on the 10 testing volumes. We compare our results with the best reported 3D methods [16,17,18,19] of the challenge, pointing the first three did not obtained those results in the challenge conditions (training on the 20 samples). Among these automatic methods ours comes in fifth position. In Table 2 we report the same validation measures (OVE: overlap error, VOD: volume difference, AVD: average distance, RMSD: root mean squared distance, MAXD: maximum distance) and inter-observer scores (see [5]) as used in the challenge. Those results have also been published online [4].

On this challenge database we get comparable results to best scored methods (see last column in Table 2). Our method also presents the advantage of being faster and requiring few samples for training. Moreover we thoroughly evaluate the robustness of liver localization in various conditions hereafter.

5.3 Evaluation on a Large Varied Database

Database Description. The database we use in this experiment is composed of 268 3D CT images coming from 127 patients with diverse medical conditions ($\sim$ 41% of patients with significant alterations of the liver shape and/or appearance). The database includes volumes with varied body shapes, fields of view (for 28% of the database the images include a large part of the trunk), resolution and use or no use of contrast agents (19% of delayed or non contrasted scans, 31% of hepatic arterial phases and 50% of portal venous phases) as shown in Fig. 5. Slices and inter-slices resolution ranges from 0.5 to 1 mm and from 0.5 to 3 mm, respectively. The 268 images have been segmented manually by an expert.

The regression forest we use in this experiment is the same one as the one used previously. In Table 3 we report the results after each step thus showing the relevance of the global-to-local strategy. For the sake of consistency we use the same evaluation measures as in the first experiment. The localization with

Table 3. Results and computational time after each step of the algorithm reported as Mean ± Standard-Deviation

	Time [sec.]	**OVE** [%]	**VOD** [%]	**AVD** [mm]	**RMSD** [mm]	**MAXD** [mm]
RF localization	1.5	-	-	10.8 ± 6.8	-	-
Proba. map	6	-	-	-	-	-
Step 1	1	57.7 ± 10.4	57.3 ± 11.1	23.8 ± 9.9	31.6 ± 12.3	84.9 ± 27.4
Step 2	10	10.4 ± 6.7	-1.0 ± 6.0	2.3 ± 3.2	4.0 ± 4.6	23.3 ± 14.9
Step 3	23	9.6 ± 6.7	-2.4 ± 5.7	1.9 ± 2.9	3.3 ± 4.1	21.5 ± 13.7
Step 4	5	8.4 ± 6.7	-1.5 ± 5.6	1.7 ± 2.4	3.0 ± 3.6	21.7 ± 14.1

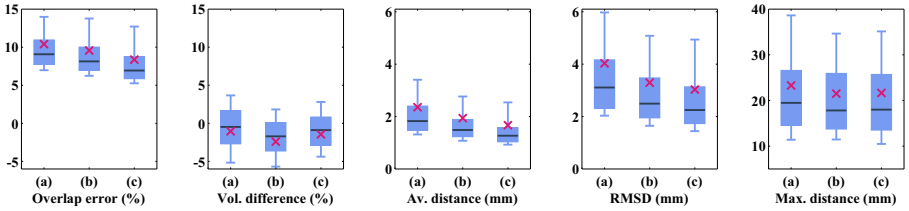


Fig. 4. Boxplots for (a) step 2 (b) step 3 and (c) step 4. They represent 1st and 9th decile, 1st and 3rd quartile, median (dark-blue dash) and mean (pink cross).

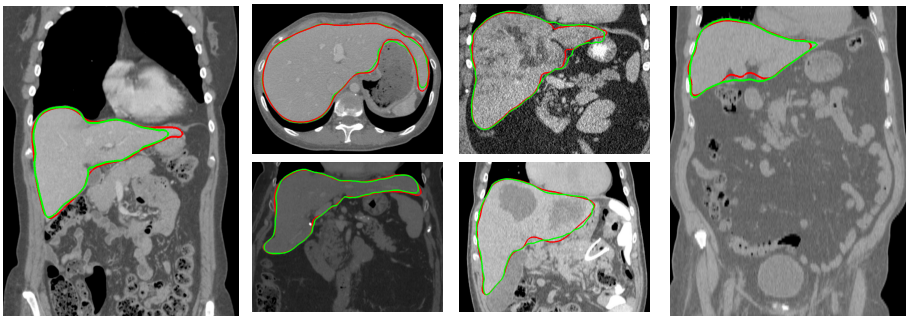


Fig. 5. Examples of segmentation results (red: ground truth, green: our result) in varied situations: different fields of view, liver contrast, shape and pathologies

regression forest is fast (1.5s) and robust as the average distance (mean distance of box faces) is of 10.8mm for a maximum of 46.8mm. In comparison the authors of [10] obtain an average distance of 15.7mm for liver localization. We again emphasize that the regression forest was trained on only 20 datasets, which further highlights the robustness of the method.

After the last refinement we obtain mean and median distances of 1.7mm and 1.3mm, respectively. The median value highlights the presence of a limited number of outliers. Indeed for more than 90% of the database the overlap error is below 15.8% and the average distance below 3mm. Outliers can be principally explained by a wrong initial position of the shape model (imprecise bounding box) and by diseases giving a very atypical appearance to the liver. In Fig. 4 we represent boxplots of the different measures, showing the relatively compact dispersion of them. These results confirm those obtained on the SLIVER07 database and are good, despite the much larger variability of the database. Fig. 5 shows the segmentation accuracy in various situations.

6 Conclusion

In this paper we proposed a fully-automatic workflow for CT liver segmentation, robust to a large variety of imaging conditions: different fields of view, different CT phases, healthy and diseased livers. Our method relies on regression forests to predict liver and heart bounding boxes, computes a probability map from an estimation of the liver intensity distribution and uses a template-based deformation algorithm to perform the liver segmentation in a global-to-local strategy.

Our state-of-the-art results on the SLIVER07 database are confirmed by an additional, extensive evaluation on a large and heterogeneous database (268 volumes). This validation demonstrates that our framework reaches an attractive balance between robustness, accuracy (mean distance to ground truth of 1.7mm) and speed (46s). We emphasize the genericity and relative simplicity of our framework, which required very limited liver-specific tuning. It is reproducible and could be improved in a number of ways. For instance, failed segmentations could be detected and easily corrected by the clinician as the template deformation framework we employ can handle user interactions [11]. Finally we believe similar workflows could be applied to other organs and imaging modalities.

References

1. Murthy, R., Nunez, R., Szklaruk, J., Erwin, W., Madoff, D.C., Gupta, S., Ahrar, K., Wallace, M.J.: Yttrium-90 microsphere therapy for hepatic malignancy: Devices, indications, technical considerations, and potential complications. Radiographics 25(suppl. 1), 41–55 (2005)
2. Campadelli, P., Casiraghi, E., Esposito, A.: Liver segmentation from computed tomography scans: A survey and a new algorithm. Artificial Intelligence in Medicine 45(2-3), 185–196 (2009)
3. Mharib, A.M., Ramli, A.R., Mashohor, S., Mahmood, R.B.: Survey on liver CT image segmentation methods. Artificial Intelligence Review 37(2), 83–95 (2011)
4. Heimann, T., Styner, M., van Ginneken, B.: Sliver07 (2007), `http://www.sliver07.org` (accessed: May 12, 2013)
5. Heimann, T., et al.: Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Transactions on Medical Imaging 28(8), 1251–1265 (2009)

6. Ling, H., Zhou, S.K., Zheng, Y., Georgescu, B., Suehling, M., Comaniciu, D.: Hierarchical, learning-based automatic liver segmentation. In: Proc. CVPR 2008, pp. 1–8 (2008)
7. Zhang, X., Tian, J., Deng, K., Wu, Y., Li, X.: Automatic liver segmentation using a statistical shape model with optimal surface detection. IEEE Transactions on Biomedical Engineering 57(10), 2622–2626 (2010)
8. 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(2), 771 (2010)
9. Cuingnet, R., Prevost, R., Lesage, D., Cohen, L.D., Mory, B., Ardon, R.: Automatic detection and segmentation of kidneys in 3D CT images using random forests. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 66–74. Springer, Heidelberg (2012)
10. Criminisi, A., Robertson, D., Konukoglu, E., Shotton, J., Pathak, S., White, S., Siddiqui, K.: Regression forests for efficient anatomy detection and localization in computed tomography scans. Medical Image Analysis (in press, 2013)
11. Mory, B., Somphone, O., Prevost, R., Ardon, R.: Real-time 3D image segmentation by user-constrained template deformation. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 561–568. Springer, Heidelberg (2012)
12. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. IEEE Transactions on Medical Imaging 29, 196–205 (2010)
13. Fitzgibbon, A.W.: Robust registration of 2D and 3D point sets. Image and Vision Computing 21(13-14), 1145–1153 (2003)
14. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Transactions on Pattern Analaysis and Machine Intelligence 14(2), 239–256 (1992)
15. He, K., Sun, J., Tang, X.: Guided image filtering. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 1–14. Springer, Heidelberg (2010)
16. Kainmüller, D., Lange, T., Lamecker, H.: Shape constrained automatic segmentation of the liver based on a heuristic intensity model. In: Proc. MICCAI Workshop 3D Segmentation in the Clinic: A Grand Challenge, pp. 109–116 (2007)
17. Wimmer, A., Soza, G., Hornegger, J.: A generic probabilistic active shape model for organ segmentation. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part II. LNCS, vol. 5762, pp. 26–33. Springer, Heidelberg (2009)
18. 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)
19. Huang, C., Jia, F., Li, Y., Zhang, X., Luo, H., Fang, C., Fan, Y.: Fully automatic liver segmentation using probability atlas registration. In: International Conference on Electronics, Communications and Control 2012, pp. 126–129 (2012)

Tumor Subtype-Specific Parameter Optimization in a Hybrid Active Surface Model for Hepatic Tumor Segmentation of 3D Liver Ultrasonograms

Myungeun Lee[1], Jong Hyo Kim[2,3], Moon Ho Park[4], Ye-Hoon Kim[4], Yeong Kyeong Seong[4], Junghoe Kim[4], Baek Hwan Cho[4], Sinsang Yu[4], and Kyoung-Gu Woo[4]

[1] Medical Research Center, Seoul National University, Seoul, Korea
melee@snu.ac.kr
[2] Department of Radiology, Seoul National University College of Medicine, Seoul, Korea
[3] Department of Transdisciplinary Studies, Graduate School of Convergence Science and Technology, Seoul National University, Suwon, Korea
kimjhyo@snu.ac.kr
[4] Data Analytics Group, Future IT Laboratories, Samsung Advanced Institute of Technology, Samsung Electronics, Yongin-Si, Korea

Abstract. Segmentation of hepatic tumors is a clinically demanding task for improving reliability in diagnosis and treatment procedures, and yet remains a challenging problem due to their highly noisy, low contrast, and blurry imaging nature. However, once correctly segmented, the shape and volume information of a tumor may provide useful information for radiological decision making. In this study, we propose an active surface model. The model combines edge, region, and contour smoothness energies. We extracted qualitative appearance features from three hepatic tumor subtypes and use them to adjust the weights of the energy terms in order to determine an optimized set of parameters for each tumor subtype. The performance of the developed method was evaluated with a dataset of 60 cases including 18 hepatic simple cysts, 18 hemangiomas, and 24 hepatocellular carcinomas, as determined by the radiologist's visual assessment. Evaluation of the results showed that our proposed method produced tumor boundaries that were equal to or better than acceptable in 87% of cases.

Keywords: Tumor segmentation, active surface model, liver ultrasound.

1 Introduction

Diagnostic ultrasound (US) is a useful clinical tool for visualizing and detecting soft tissue lesions such as hepatic tumors without any deteriorating effects [1]. Segmentation of hepatic tumors is a clinically demanding task for improving reliability in diagnosis, treatment planning, and monitoring of treatment follow-up. Once correctly segmented, the shape and volume information of a tumor may provide important information in radiological decision making. However, robust algorithms for obtaining correct segmentation of US images rare, particularly for hepatic tumors. The high

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 251–258, 2013.

level of irregular speckle patterns and artifacts present in US images complicates most associated computer algorithms.

Region growing and active contour approaches are the two most widely studied methods used for tumor segmentation. The region growing method [2] first requires the user to identify a seed point and grow the boundary as long as the homogeneity of the region is maintained using geometric priority and feature vector space of the seed point as criteria, in which the interference of speckle noise makes the classification process highly unreliable. The active contour model [3] uses an internal energy model to deform the initial contour toward the desired object boundary. The gradient and contour curvature irregularities observed in US images disturb the contour evolution process in active contour models and frequently lead to undesirable termination of segmentation.

To overcome these drawbacks, we propose an active surface model that enables robust segmentation of tumor boundaries in three dimensional (3D) liver US images by considering tumor subtype-specific appearance features in an internal energy model. In particular, the proposed method employs a hybrid speed function that comprises three energy terms such as edge, region, and smoothing. The edge term drives the model toward the object boundary by calculating the inner product of the surface normal and the gradient of the input image. The region term searches for a dividing boundary that minimizes the intra-regional variances in pixel intensity distribution. The smoothing term regularizes the boundary contour to avoid unnecessary spurious variation caused by speckle noise. We adapted this hybrid active surface model to segment reliably hepatic tumors in 3D US by obtaining sets of weighting factors for the three energy terms. The factors were optimized for the three tumor subtypes most often encountered in liver US.

2 Materials and Methods

2.1 Materials

Our method has been developed and tested with a set of 60 volumetric liver ultrasound images each containing at least one tumor acquired at department of radiology, SMC (Samsung Medical Center). The 60 cases were classified into 3 categories according to the tumor subtypes; 18 hepatic simple cysts**,** 18 hemangiomas, and 24 hepatocellular carcinoma (HCC). The classification was made by an expert radiologist with more than 10 years of experience.

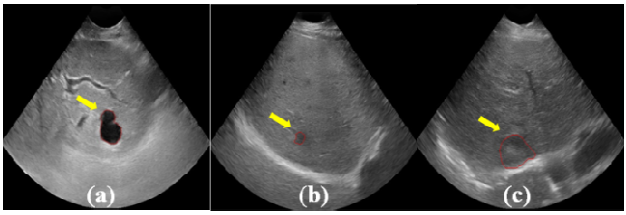


Fig. 1. Example ultrasound images showing three tumor subtypes: (a) cyst, (b) hemangioma, and (c) HCC. Each volumetric image contained a slice image with tumor boundaries arrow drawn by a radiologist.

The images were obtained with a 3D ultrasound probe on a diagnostic ultrasound equipment (X6-1, Philips). Figure 1 shows example images of three subtypes of hepatic tumor and includes tumor boundaries manually drawn by a radiologist. In general, cyst has strong contrast with much darker pixel value than that of parenchymal tissue. On the other hand, the hemangioma and HCC have relatively round shapes with ambiguous contrast and fuzzy boundary edges.

Each volumetric image consisted of 256 slices. In addition, for each case the radiologist provided a manually drawn tumor contour on a selected slice image which best depicts the tumor for each case. The contour containing images were used to fine tune the weighting factors of our active surface model with the tuning based on the appearance pattern of each of the tumor subtypes. After developing and adapting the proposed technique, the model marked the tumor contours as a color overlay on the slice images. The color overlay contours were then sent to the radiologist for evaluation of the model's segmentation results.

2.2 Methods: Active Surface Model with Hybrid Speed Function

The level set approach was introduced by S.Osher and J.A.Sethian [3]. They modeled the propagating curve as a specific level set of a higher dimensional surface with promising results [4]. A main drawback of this approach in application to segmenting US images is that the method only fits to an image including regions of piecewise constant intensities, which often does not fit to the liver US image resulting in incorrect identification of boundaries.

In this study, we propose an active surface model using a hybrid speed function which consists of three energy terms such as edge, region, and the smoothing term. The edge term is for driving the contour toward the tumor edges where the tumor has distinct boundary. However, in case the tumor does not have distinct boundary, the insufficient magnitude of gradient fails to stop the level set evolution at the desirable tumor boundary, which makes leakage often inevitable. Hence as an alternative, we employ the gradient vector flow (GVF) field proposed in Xu et al. [5] as a edge indicator vector embedded with a speed term. The particular advantage of the GVF includes its insensitivity to initialization and its ability to move into boundary concavities. In addition, as the GVF provides a smooth field directing towards strong edge, it represents a robust edge measure less sensitive to noise and speckle patterns. Thus, we obtain a speed function, which is utilized as a robust edge term as follows:

$$F = sign(< V_{GVF}, \nabla \phi >) div(V_{GVF}). \quad (1)$$

where $\mathbf{V}_{GVF}$ is a GVF field.

For the second energy term, we use the region information on target. The role of the region term is to define a region boundary based on a probability model for intensity distribution on target and background regions assuming that each region has distinct probability function. Use of region information in segmentation is known to overcome some of the weaknesses in the boundary-based model such as dependency of local information and initialization by optimally partitioning a given image into

multiple homogenous regions. The joint probability of intensity values I observed at a given image consisting of two homogeneous partitions $\Omega = \{\Omega_1, \Omega_2\}$ is given by

$$\begin{aligned} p(I \mid \{\Omega_1, \Omega_2\}) &= p(I \mid \Omega_1) p(I \mid \Omega_2) \\ &= \prod_{w \in \Omega_1} p_1(I(w)) \prod_{w \in \Omega_2} p_2(I(w)) \end{aligned}. \tag{2}$$

In this case, the optimal segmentation is found by minimizing the energy functional which takes logarithm of probability functions. By introducing the Heaviside function H and after a standard rescaling that involves replacing $\delta(\phi)$ with $|\nabla\phi|$, we have the following equation:

$$\phi_t = |\nabla\phi(w)| (\log p_1(I(w)) - \log p_2(I(w))) . \tag{3}$$

Here, we have assumed that the probability distribution of the intensity value on each region takes Gaussian distribution.

The smoothing term regularizes the surface curvature so as to make smooth surface preventing irregular fluctuation caused by image noise and speckle patterns. The curvature term used in geodesic active surface[7][8][9] is known to serve as a good regularizer in a noisy image. Solving the Euler-Lagrange equation gives the following evolution equation

$$\frac{\partial \phi}{\partial t} = \left(g(I) div \left(\frac{\nabla \phi}{|\nabla \phi|}\right) - < \nabla g(I), \frac{\nabla \phi}{|\nabla \phi|} > \right) |\nabla \phi| . \tag{4}$$

By combining three terms with appropriate weights, we obtain a speed function of the evolving surface

$$\begin{aligned} F(C) &= \alpha (sign(< \mathbf{V}_{GVF}, \nabla \phi >) div(\mathbf{V}_{GVF})) + \beta(\log p_1(I(w)) \\ &- \log p_2(I(w))) + \gamma \left(g(I) div\left(\frac{\nabla \phi}{|\nabla \phi|}\right) - < \nabla g, \frac{\nabla \phi}{|\nabla \phi|} > \right) \end{aligned} \tag{5}$$

where $\alpha, \beta, and\ \gamma$ are weighting factor, allowing the user to control the behavior of the active surface.

Finally, the update equation is given by the following solution:

$$\phi^{(t)} = \phi^{(t-1)} + \Delta t \cdot F \cdot |\nabla \phi^{(t-1)}| . \tag{6}$$

where Δt represents discrete time steps and F is a speed term.

3 Experimental Results

Table 1 shows four appearance features and associated scores that were obtained through consensus of an experts panel for the three hepatic tumor subtypes considered in this study. Although result of qualitative visual scoring, these data were regarded as essential information for use in linking the experienced human visual system results with the computer algorithm system results in our study.

Table 1. Appearance features and scores for three hepatic tumor subtypes

Tumor Subtype / Feature	CYST	HEMANGIOMA	HCC
Tumor Edge	+ +	–	–
Roundness	+	+ +	+
Contrast	+ +	+	–
Internal Texture	–	+ +	+ +
++: high, +: mid, -: low			

The information in Table 1 was used to adjust the weights of speed function so as to determine a fine tuned set of weights best fitting to each tumor subtype. Edge and internal texture information were used to adjust the weight for edge term, roundness for smoothing term, add contrast for region term.

Figure 2 shows the way initial surface was given in our study. First, we set a $point_1$ around the center of the tumor, and the $point_2$ to mark a point outside the tumor. Then, a circle was determined to have a radius with half the distance between $point_1$ and $point_2$, which was given to the active surface model as an initial surface.

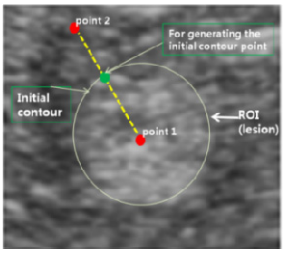


Fig. 2. Example image showing the approach used to determine the initial surface.

Figure 3 shows intermediate US segmentation results showing the change of behavior according to different combinations of speed parameters for three hepatic tumor subtypes.

A parameter set with a strong edge term restricted expansion of boundary contour evolution resulting under segmentation in Fig. 3(b). A strong region term produced boundary curves with highly irregular jags while the strong smoothing term controlled smoothness of the segmented tumor boundary in Figs. 3(c) and 3(d). After performing a combinatorial search for an optimal parameter set with visual evaluation of the results, we determined three sets of optimal parameters that were fine tuned for the three tumor subtypes: cyst set (β =0.1, β =0.8, and β =0.1), hemangioma set (β =0.1, β =0.6, and β =0.3), and HCC set (β =0.1, β =0.7, and β =0.2).

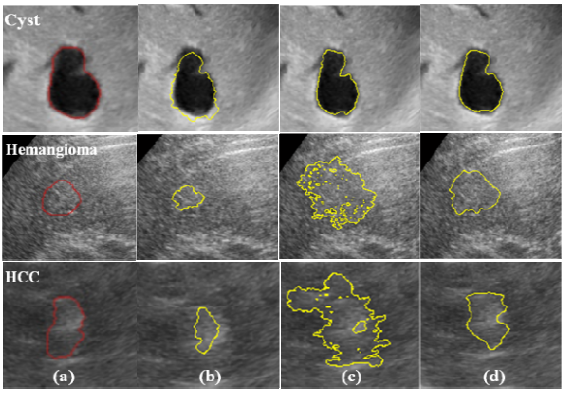


Fig. 3. Intermediate segmentation results showing the behavior of the proposed active surface model according to three control parameter settings: (a) reference, (b) a set with a strong edge term (α =0.8, α =0.1, α =0.1), (c) set with a strong region term(α =0.1, α =0.8, α =0.1), (d) strong smoothing term(α =0.1, α =0.1, α =0.8)

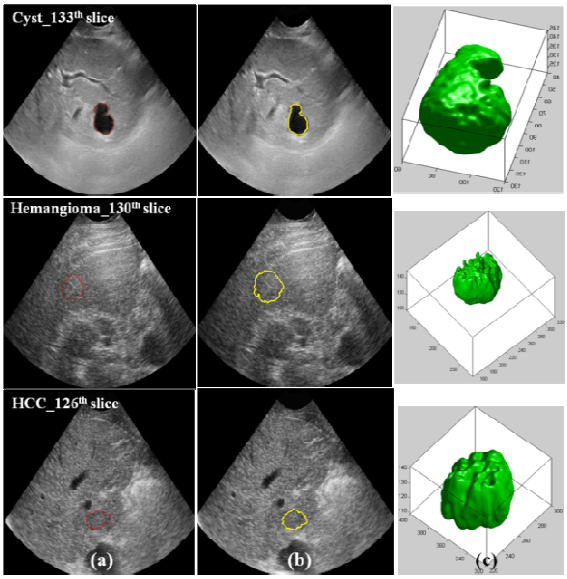


Fig. 4. Example of segmentation results from an optimized parameter set applied to each of the three hepatic tumor subtypes: (a) reference contours, (b) contours obtain from the proposed method, and (c) volume rendered segmentation results

Table 2 presents quantitative evaluation results of the segmentation output as rated by the radiologist. Rating scores 5 to 1 represented accurate, nearly accurate, acceptable, unacceptable, and inappropriate, respectively. The evaluation result shows that 87% of the segmentation results obtained by using our proposed method were equal to or better than acceptable when rated by an expert radiologist.

Table 2. Quantitative evaluation result for the segmentation output by the proposed method

Grade / Tumor subtype	1	2	3	4	5	Total (# of cases)
CYST	0.0%	5.6%	5.6%	44.4%	44.4%	18
HEMANGIOMA	5.6%	11.1%	27.8%	33.3%	22.2%	18
HCC	8.3%	8.3%	29.2%	37.5%	16.7%	24
Average (%)	5.0%	8.3%	21.7%	38.3%	26.7%	60 / 100%

In addition, we compared US segmentation results from the proposed method and conventional level set method following to Chan-Vese model [6] for a small subset of cases. Fig. 5 (a)-(c) compares the three segmented results for typical tumor subtypes. For the cyst, the three contours are nearly identical whereas for the hemangioma and HCC, our proposed method agrees better than that of conventional level set method method.

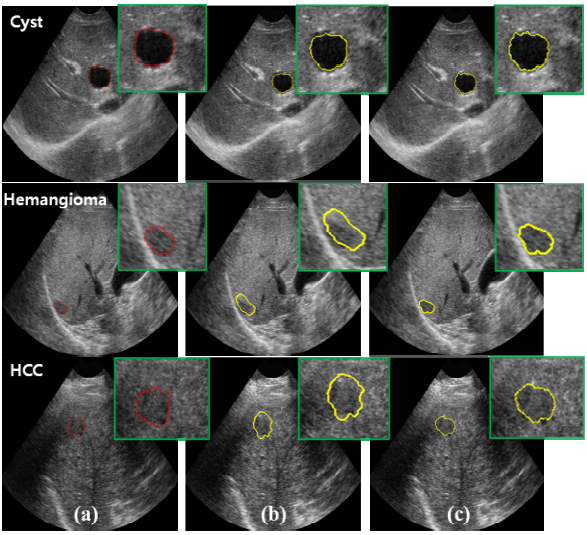


Fig. 5. Comparison of the segmentation results for each tumor subtype from (a) reference, (b) Chan-Vese model, and (c) our proposed method

4 Conclusion

This study proposed a supervised optimization procedure for tumor US segmentation algorithm based on active surface model combining edge and region information of ultrasonogram within and surrounding the tumor. We could derive an optimal set of

parameters tuned to three different hepatic tumor subtypes, which produced reasonably acceptable segmentation results. Our scheme has a potential to be used in tumor volumetric from 3D US examinations of liver. In addition, the method can be used for quantitative analysis of tumor characteristics for CAD applications.

References

1. Suganya, R., Rajaram, S.: Content Based Image Retrieval of Ultrasound Liver Diseases Based on Hybrid Approach. American Journal of Applied Sciences 9(6), 938–945 (2012)
2. Hao, X., Bruce, C., Pislaru, C., et al.: A novel region growing method for segmenting ultrasound images. In: IEEE Ultrasonic Symposium, vol. 2, pp. 1717–1720 (2000)
3. Osher, S.J., Sethian, J.A.: Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. Journal of Computer Physics 79, 12–49 (1998)
4. Osher, S., Paragios, N.: Geometric level set methods in imaging, vision, and graphics. Springer, Heidelberg (2003)
5. Xu, C., Prince, J.L.: Snakes, Shapes, and Gradient Vector Flow. IEEE Transacts on Image Processing 7(3), 359–369 (1998)
6. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. on Image Processing 10(2), 266–277 (2001)
7. Kimmel, R.: Fast edge integration. In: Geometric Level Set Methods in Imaging Vision, and Graphics, pp. 59–79. Springer (2006)
8. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Int. J. Comput. Vision 22(1), 61–79 (1997)
9. Osher, S., Paragios, N.: Geometric Level Set Methods in Imaging Vision, and Graphics. Springer, Heidelberg (2003)

Continuous-Time Flow-Limited Modeling by Convolution Area Property and Differentiation Product Rule in 4-Phase Liver Dynamic Contrast-Enhanced CT

Sang Ho Lee[1], Yasuji Ryu[1], Koichi Hayano[2], and Hiroyuki Yoshida[1]

[1] 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, yryu@partners.org
[2] Division of Abdominal Imaging and Intervention, Department of Radiology,
Massachusetts General Hospital, 55 Fruit St., Boston, Massachusetts 02114, USA
khayano@partners.org

Abstract. Parametric-fitting approaches for tracer kinetic modeling depend on the capability of a computational method to describe underlying physiologic processes that cause temporal intensity changes in dynamic contrast-enhanced (DCE) images. Rapid scan techniques allow perfusion CT imaging with high temporal resolution. In clinical practice, however, the perfusion CT protocol is especially a trade-off between the number of data points and the total radiation dose. Considering availability and radiation exposure, use of DCE-CT imaging derived from 4 temporal phases, which include precontrast, arterial, portal, and delayed phases, is highly desirable for the liver. However, low-temporal-resolution images like 4-phase liver DCE-CT present several barriers to modeling of tracer kinetics because of the lack of temporal enhancement data, which limits obtaining reliable physiologic information. The major reason for the limited application of a tracer kinetic model in temporally sparse dynamic data is that general computational algorithms such as deconvolution techniques require discretizing of arterial (or portal-vein) and tissue curves for estimation of kinetic parameters, leading to an unstable computational solution. The numerical instability due to the discretization of the enhancement curves can be more pronounced in the low-temporal-resolution data like those gleaned from 4-phase DCE-CT. For this reason, we propose a novel dual-input continuous-time tracer kinetic modeling method based on a new mathematical approach that uses the convolution area property and the differentiation product rule, without any discretization of the enhancement curves. This model was applied to case studies of hepatocellular carcinoma in 4-phase DCE-CT to illustrate the potential effectiveness of continuous-time tracer kinetic modeling. The proposed analytic scheme was shown to be feasible for estimation of kinetic parameters even in 4-phase liver DCE-CT, potentially being a practical guide for tracer kinetic model-based curve-fitting in temporally sparse data.

Keywords: Continuous-time tracer kinetic modeling, convolution area property, differentiation product rule, four-phase dynamic contrast-enhanced CT.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 259–269, 2013.

1 Introduction

Dynamic contrast-enhanced CT (DCE-CT) is a noninvasive in-vivo technique that provides valuable information about tissue and tumor microvascular function [1]. The technique measures the vascular support of tumors through analysis of temporal changes during sequential imaging, involving the intravenous injection of an iodinated contrast agent (CA). The temporal changes in CT attenuation can be analyzed by fitting of a tracer kinetic model to the sampled tracer enhancement curves to yield quantitative estimates of kinetic parameters associated with the tissue microcirculation [2,3]. This approach depends on the capability of a computational method to describe underlying physiologic processes that cause temporal intensity changes in DCE images [4,5].

Recent technologic advances and rapid scan techniques have improved the applicability of perfusion CT (PCT) imaging with high temporal resolution. Whole-liver PCT imaging has the potential to provide both high-temporal-resolution and high-spatial-resolution imaging of the entire liver for the detection of hepatic tumors. In clinical practice, however, the PCT protocol is a trade-off between the amount of data points collected and the total radiation dose [6,7], and the PCT evaluation, as such, is usually limited to selected portions of the liver [8]. Therefore, the clinical implementation of whole-liver PCT imaging requires the exploration of methods for reducing the radiation exposure [8].

The desire for increased spatial coverage and reduced radiation dose has resulted in various PCT protocols [9]. Considering the availability of CT scans and the radiation exposure, use of DCE-CT imaging derived from 4 temporal phases, which include pre-contrast, arterial, portal, and delayed phases, is a clinically favored option for whole-liver imaging [10,11]. However, the low-temporal-resolution images present several barriers to modeling of tracer kinetics because of the lack of temporal enhancement data, which limits obtaining reliable physiologic information. Three major issues hamper the establishment of a model-generated solution for tracer kinetics in 4-phase liver DCE-CT: 1) absence of temporal enhancement information during the first pass (early arterial phase), 2) invisibility of time lag to account for the difference in bolus arrival times between the hepatic artery (or portal vein) and liver tissue, and 3) numerical instability of deconvolution operation because it frequently yields non-physiologic oscillations (aliasing or noise) in the computation of the solution for the tissue residue function [4].

The major reason for these difficulties is that conventional computational algorithms such as deconvolution techniques require discretizing arterial (or portal vein) and tissue curves for estimation of kinetic parameters, leading to a computationally unstable solution. This numerical instability due to the discretization of the enhancement curves can be more pronounced in the low-temporal-resolution data in DCE-CT. To solve this problem, we propose a novel dual-input continuous-time tracer kinetic modeling method, which is derived under a flow-limited condition, based on a new mathematical approach that uses the convolution area property and the differentiation product rule, without any discretization of the enhancement curves. Our purpose in this study was to illustrate the potential effectiveness of the proposed analytic method through case studies of hepatocellular carcinoma (HCC) in 4-phase liver DCE-CT.

2 Methods

2.1 Flow-Limited Tissue Compartment Model

We consider the CA blood concentration-time curves in the hepatic artery and portal vein (PV), $C_{\mathrm{A}}(t)$ and $C_{\mathrm{PV}}(t)$ (in g/ml), and two sources of total hepatic plasma flow F (in ml/min), i.e.: arterial and portal plasma flows, F_{A} and F_{PV}, respectively. In the flow-limited condition of the two compartment exchange (2CX) model [3], the administered CA between the plasma and interstitial compartment is fast when compared to the plasma flow so that the distribution space of the CA in the tissue can be approximated by a single tissue compartment with the relative distribution volume v_{D}, i.e.: $v_{\mathrm{D}} = v_{\mathrm{P}} + v_{\mathrm{I}}$ [3], where v_{P} and v_{I} are the volume fractions of the plasma and interstitial space within the tissue volume. Thus, the mass balance equation in the flow-limited condition of the 2CX model can be described as the follows:

$$v_{\mathrm{D}}\frac{dC_{\mathrm{D}}(t)}{dt} = \frac{dC_{\mathrm{T}}(t)}{dt} = \frac{F}{V_{\mathrm{D}}}\left[\frac{v_{\mathrm{D}}C_{\mathrm{in}}(t)}{1-H_{\mathrm{LV}}} - C_{\mathrm{T}}(t)\right], \tag{1}$$

where $C_{\mathrm{D}}(t) = C_{\mathrm{T}}(t)/v_{\mathrm{D}}$ is the concentration within the tissue distribution space, $V_{\mathrm{D}} = v_{\mathrm{D}}V_{\mathrm{T}}$ is the distribution volume, and H_{LV} is the hematocrit of blood in the large vessels ($\cong 0.45$) [3]. $C_{\mathrm{in}}(t) = \gamma C_{\mathrm{A}}(t) + (1-\gamma)C_{\mathrm{PV}}(t)$ is the net input blood concentration, where $\gamma = F_{\mathrm{A}}/F$ is the arterial flow fraction. The analytic solution for $C_{\mathrm{T}}(t)$ becomes

$$C_{\mathrm{T}}(t) = \frac{F}{V_{\mathrm{T}}}R_{\mathrm{T}}(t)\otimes\frac{C_{\mathrm{in}}(t)}{1-H_{\mathrm{LV}}} = \frac{F}{V_{\mathrm{T}}}e^{-\frac{1}{v_{\mathrm{D}}V_{\mathrm{T}}}t}\otimes\frac{C_{\mathrm{in}}(t)}{1-H_{\mathrm{LV}}}, \tag{2}$$

where F/V_{T} is the total hepatic perfusion, $R_{\mathrm{T}}(t) = e^{-\frac{1}{v_{\mathrm{D}}V_{\mathrm{T}}}t}$ is the tissue's impulse residue response function, and $\otimes$ denotes the convolution operator.

2.2 Arterial Input Function Model

We modeled a full arterial input function (AIF), $C_{\mathrm{A}}(t)$, including the first-pass (early arterial phase) and recirculating inputs to describe invisible temporal enhancements in the intervals of 4-phase DCE-CT data. The $C_{\mathrm{A}}(t)$ was decomposed into a bolus model denoted as $C_{\mathrm{B}}(t)$ and a body transfer function (BTF) denoted as $G(t)$ [12]. These two elements may take any form, but restricting them to suitable mathematical functions guarantees that the calculations can be implemented analytically. Even with these constraints, a realistic full AIF model can be defined that contains a first-pass peak, a recirculation peak, and the later temporal phases. The $C_{\mathrm{A}}(t)$ can be a superposition of the bolus shape and its shape after modification by the BTF. By considering the time lag of the bolus arrival to the artery, $t_{\mathrm{Lag,A}}$,

$$C_{\mathrm{A}}(t) = \begin{cases} 0 & t \le t_{\mathrm{Lag,A}} \\ C_{\mathrm{B}}(t - t_{\mathrm{Lag,A}}) + C_{\mathrm{B}}(t - t_{\mathrm{Lag,A}})\otimes G(t) & t > t_{\mathrm{Lag,A}}. \end{cases} \tag{3}$$

Note that the recirculating input delay was assumed to be negligible for simplicity. We used $C_{\mathrm{B}}(t) = a_{\mathrm{B}} t e^{-\mu_{\mathrm{B}} t}$ for the bolus model, and $G(t) = a_{\mathrm{G}} e^{-\mu_{\mathrm{G}} t}$ for the BTF. Performing the forward and inverse Laplace transforms of the equation (3) leads to

$$C_{\mathrm{A}}(t) = \begin{cases} 0 & t \le t_{\mathrm{Lag,A}} \\ A_{\mathrm{B}}\left(t - t_{\mathrm{Lag,A}}\right) e^{-\mu_{\mathrm{B}}\left(t - t_{\mathrm{Lag,A}}\right)} + A_{\mathrm{G}}\left(e^{-\mu_{\mathrm{G}} t} - e^{-\mu_{\mathrm{B}}\left(t - t_{\mathrm{Lag,A}}\right)}\right) & t > t_{\mathrm{Lag,A}}, \end{cases} \tag{4}$$

with $A_{\mathrm{B}} = a_{\mathrm{B}} - a_{\mathrm{B}} a_{\mathrm{G}}/(\mu_{\mathrm{B}} - \mu_{\mathrm{G}})$ and $A_{\mathrm{G}} = a_{\mathrm{B}} a_{\mathrm{G}}/(\mu_{\mathrm{B}} - \mu_{\mathrm{G}})^2$.

2.3 Portal Venous Input Function Model

Generally, in the systemic circulation of CA, it has been known that the PV enhancement results from output via both the spleen and the intestine from the abdominal aorta [13]. However, any tracer kinetic model that considers the spleen and intestine compartments causes difficulty in fitting 4-phase DCE-CT data due to the higher number of kinetic parameters contained in the model than the actual number of available temporal data. In this situation, using the simplest possible model that fits the data is desirable, ensuring that the fit is reliable. Thus, we modulated a PV input function (PVIF) by using the Tofts-Kety (TK) model [14]. Because the contrast present in the PV comes from the abdominal aorta, the PVIF was modeled considering the AIF as a single input for the PV enhancement, i.e.:

$$C_{\mathrm{PV}}(t) = \begin{cases} 0 & t \le t_{\mathrm{Lag,A}} + t_{\mathrm{Lag,PV}} \\ K_{\mathrm{PV}}^{\mathrm{Trans}} e^{-k_{\mathrm{PV}} t} \otimes C_{\mathrm{A}}\left(t - t_{\mathrm{Lag,PV}}\right) & t > t_{\mathrm{Lag,A}} + t_{\mathrm{Lag,PV}}, \end{cases} \tag{5}$$

where $t_{\mathrm{Lag,PV}}$ denotes the time lag of the bolus arrival from the hepatic artery to the PV. The $K_{\mathrm{PV}}^{\mathrm{Trans}}$ is related to the blood flow between the artery and PV. In parallel, the higher the concentration in the PV, the more this contrast will tend to diffuse into the interstitial space and to the other blood plasma. The k_{PV} is then connected to the speed that characterizes this diffusion process. From equation (5), we derived the solution of the continuous-time TK model by using the convolution area property and the differentiation product rule so as not to discretize the arterial and PV enhancement curves. The area under the convolution is the product of areas under the factors, i.e.: for $t > t_{\mathrm{Lag,A}} + t_{\mathrm{Lag,PV}}$,

$$\int_0^t C_{\mathrm{PV}}(\tau) d\tau = K_{\mathrm{PV}}^{\mathrm{Trans}} \left[\int_0^t e^{-k_{\mathrm{PV}} \tau} d\tau\right] \cdot \left[\int_{t_{\mathrm{Lag,A}}}^{t - t_{\mathrm{Lag,PV}}} C_{\mathrm{A}}(\tau) d\tau\right], \tag{6}$$

where we defined $R_{\mathrm{PV}}(t) = e^{-k_{\mathrm{PV}} t}$ as the PV residue function. For calculating $C_{\mathrm{PV}}(t)$, this equation requires differentiation with use of the product rule, which is a formal rule for differentiating problems where one function is multiplied by another, i.e.: for $t > t_{\mathrm{Lag,A}} + t_{\mathrm{Lag,PV}}$,

$$C_{\mathrm{PV}}(t) = K_{\mathrm{PV}}^{\mathrm{Trans}}\left[C_{\mathrm{A}}\left(t - t_{\mathrm{Lag,PV}}\right)\int_{0}^{t} R_{\mathrm{PV}}(\tau)d\tau + R_{\mathrm{PV}}(t)\int_{t_{\mathrm{Lag,A}}}^{t-t_{\mathrm{Lag,PV}}} C_{\mathrm{A}}(\tau)d\tau\right] \quad (7)$$

$$= K_{\mathrm{PV}}^{\mathrm{Trans}}\left[C_{\mathrm{A}}\left(t - t_{\mathrm{Lag,PV}}\right)\frac{1 - R_{\mathrm{PV}}(t)}{k_{\mathrm{PV}}} + R_{\mathrm{PV}}(t)I_{\mathrm{A}}\left(t - t_{\mathrm{Lag,PV}}\right)\right],$$

where $I_{\mathrm{A}}(t) = \frac{A_{\mathrm{B}}}{{\mu_{\mathrm{B}}}^2}\left[1 - \left\{1 + \mu_{\mathrm{B}}\left(t - t_{\mathrm{Lag,A}}\right)\right\}e^{-\mu_{\mathrm{B}}\left(t - t_{\mathrm{Lag,A}}\right)}\right] + A_{\mathrm{G}}\left[\frac{e^{-\mu_{\mathrm{G}}t_{\mathrm{Lag,A}}} - e^{-\mu_{\mathrm{G}}t}}{\mu_{\mathrm{G}}} - \frac{1 - e^{-\mu_{\mathrm{B}}\left(t - t_{\mathrm{Lag,A}}\right)}}{\mu_{\mathrm{B}}}\right]$. The resultant fitting AIF and PVIF curves are shown in Fig. 1. The fitting AIF curve demonstrates a full-pass enhancement including the first-pass and recirculation inputs, and the fitting PVIF curve reveals a relatively more complex pattern modified by the $R_{\mathrm{PV}}(t)$ and $t_{\mathrm{Lag,PV}}$.

2.4 Liver Parameter Calculation

An approximate measure of the contrast that enters the liver by the hepatic artery and PV enables modeling of liver kinetic parameters. Like the PVIF model, we applied the simplest possible model in the flow-limited condition of the 2CX model for calculation of liver kinetic parameters. However, there still remains a challenging task in terms of data-fitting and parameter precision. Considering a bolus of CA injected into the liver tissue, we can define the CA concentration in the tissue, $C_{\mathrm{T}}(t)$, in terms of two functions. The first is the tissue residue function, $R_{\mathrm{T}}(t)$, which represents a fraction of the CA that is still present in the region of interest (ROI) at time t following an ideal instantaneous bolus injection. The second is the input blood function $C_{\mathrm{in}}(t)$. The observed tissue curve represents a combination of the effects of the input function and the inherent tissue properties, which can be represented as the total hepatic perfusion, $\frac{F}{V_{\mathrm{T}}}$ multiplied by the input function with $R_{\mathrm{T}}(t)$, as shown in the equation (2). Generally, to fit the model, the effect of the input function on the tissue curve is removed by use of a mathematical process known as deconvolution for derivation of $R_{\mathrm{T}}(t)$. However, deconvolution is an unstable computational solution [15], and this numerical instability can be intensified when deconvolution is applied to low-temporal-resolution data like those gleaned from 4-phase liver DCE-CT [6]. The main reason is that the first step in the development of the deconvolution algorithm is to discretize $R_{\mathrm{T}}(t)$ and $C_{\mathrm{in}}(t)$, i.e.:

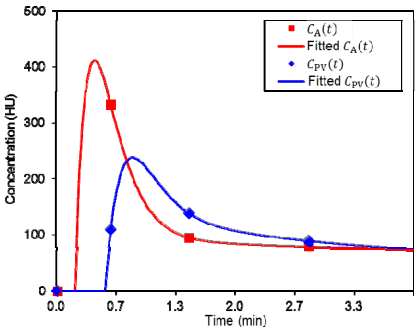


Fig. 1. Example of fitting the AIF and PVIF by the proposed models in the 4-phase DCE-CT

$C_{\mathrm{T}}(t_k) = \frac{1}{1-H_{\mathrm{LV}}}\frac{F}{V_{\mathrm{T}}}\sum_{i=1}^{k} R_{\mathrm{T}}(t_i)C_{\mathrm{in}}(t_k - t_i)\,\Delta\, t_i + \varepsilon$. Thus, the more temporally sparse the dynamic information is, the greater the error ε would be. To solve this problem, we derived a novel dual-input continuous-time tracer kinetic model that describes the time lag of the bolus arrival from the input blood to the liver tissue ($t_{\mathrm{Lag,T}}$) by applying the convolution area property and the differentiation product rule like the PVIF model between $R_{\mathrm{T}}(t)$ and $C_{\mathrm{in}}(t)$. The $t_{\mathrm{Lag,T}}$ can be imposed on $C_{\mathrm{in}}(t)$ in calculation of $C_{\mathrm{T}}(t)$, i.e.: $C_{\mathrm{T}}(t) = \frac{F}{V_{\mathrm{T}}}R_{\mathrm{T}}(t)\otimes\frac{C_{\mathrm{in}}(t-t_{\mathrm{Lag,T}})}{1-H_{\mathrm{LV}}}$. We decoupled the arterial contribution to $C_{\mathrm{T}}(t)$, denoted as $C_{\mathrm{T,A}}(t)$, and the PV contribution to $C_{\mathrm{T}}(t)$, denoted as $C_{\mathrm{T,PV}}(t)$, respectively, so that $C_{\mathrm{T}}(t) = C_{\mathrm{T,A}}(t) + C_{\mathrm{T,PV}}(t)$, where $C_{\mathrm{T,A}}(t) = \frac{F_{\mathrm{A}}}{V_{\mathrm{T}}}R_{\mathrm{T}}(t)\otimes\frac{C_{\mathrm{A}}(t-t_{\mathrm{Lag,T}})}{1-H_{\mathrm{LV}}}$ and $C_{\mathrm{T,PV}}(t) = \frac{F_{\mathrm{PV}}}{V_{\mathrm{T}}}R_{\mathrm{T}}(t-t_{\mathrm{Lag,T}})\otimes\frac{C_{\mathrm{PV}}(t)}{1-H_{\mathrm{LV}}} = \frac{F_{\mathrm{PV}}}{V_{\mathrm{T}}}K_{\mathrm{PV}}^{\mathrm{Trans}}R_{\mathrm{T}}(t)$ $\otimes R_{\mathrm{PV}}(t)\otimes\frac{C_{\mathrm{A}}(t-t_{\mathrm{Lag,PV}}-t_{\mathrm{Lag,T}})}{1-H_{\mathrm{LV}}}$. Likewise, the convolution area property can be applied to $C_{\mathrm{T,A}}(t)$ and $C_{\mathrm{T,PV}}(t)$ with $t_{\mathrm{Lag,T}}$, i.e.:

$$\int_0^t C_{\mathrm{T,A}}(\tau)d\tau = \frac{F_{\mathrm{A}}}{V_{\mathrm{T}}}\left[\int_0^t R_{\mathrm{T}}(\tau)d\tau\right]\cdot\left[\int_{t_{\mathrm{Lag,A}}}^{t-t_{\mathrm{Lag,T}}}\frac{C_{\mathrm{A}}(\tau)}{1-H_{\mathrm{LV}}}d\tau\right] \tag{8}$$

$$\int_0^t C_{\mathrm{T,PV}}(\tau)d\tau = \frac{F_{\mathrm{PV}}}{V_{\mathrm{T}}}K_{\mathrm{PV}}^{\mathrm{Trans}}\left[\int_0^t R_{\mathrm{T}}(\tau)d\tau\right]\cdot\left[\int_0^t R_{\mathrm{PV}}(\tau)d\tau\right]\cdot\left[\int_{t_{\mathrm{Lag,A}}}^{t-t_{\mathrm{Lag,PV}}-t_{\mathrm{Lag,T}}}\frac{C_{\mathrm{A}}(\tau)}{1-H_{\mathrm{LV}}}d\tau\right]. \tag{9}$$

Note that $C_{\mathrm{T,A}}(t)$ and $C_{\mathrm{T,PV}}(t)$ are defined only for $t > t_{\mathrm{Lag,A}} + t_{\mathrm{Lag,T}}$ and $t > t_{\mathrm{Lag,A}} + t_{\mathrm{Lag,PV}} + t_{\mathrm{Lag,T}}$, respectively. Otherwise, $C_{\mathrm{T,A}}(t) = C_{\mathrm{T,PV}}(t) = 0$. By defining $R_{\mathrm{T,PV}}(t) = R_{\mathrm{T}}(t)\otimes R_{\mathrm{PV}}(t)$ and applying the differentiation product rule to the equations (8) and (9), $C_{\mathrm{T,A}}(t)$ and $C_{\mathrm{T,PV}}(t)$ become

$$\begin{aligned} C_{\mathrm{T,A}}(t) &= \frac{F_{\mathrm{A}}}{V_{\mathrm{T}}}\left[\frac{C_{\mathrm{A}}(t-t_{\mathrm{Lag,T}})}{1-H_{\mathrm{LV}}}\int_0^t R_{\mathrm{T}}(\tau)d\tau + R_{\mathrm{T}}(t)\int_{t_{\mathrm{Lag,A}}}^{t-t_{\mathrm{Lag,T}}}\frac{C_{\mathrm{A}}(\tau)}{1-H_{\mathrm{LV}}}d\tau\right] \\ &= \frac{\gamma}{1-H_{\mathrm{LV}}}\left[v_{\mathrm{D}}\{1-R_{\mathrm{T}}(t)\}C_{\mathrm{A}}(t-t_{\mathrm{Lag,T}}) + \frac{F}{V_{\mathrm{T}}}R_{\mathrm{T}}(t)I_{\mathrm{A}}(t-t_{\mathrm{Lag,T}})\right] \end{aligned} \tag{10}$$

$$C_{\mathrm{T,PV}}(t) = \frac{F_{\mathrm{PV}}}{V_{\mathrm{T}}}\left[\frac{C_{\mathrm{A}}\left(t - t_{\mathrm{Lag,PV}} - t_{\mathrm{Lag,T}}\right)}{1 - H_{\mathrm{LV}}}\int_{0}^{t} R_{\mathrm{T,PV}}(\tau)d\tau + R_{\mathrm{T,PV}}(t)\int_{t_{\mathrm{Lag,A}}}^{t - t_{\mathrm{Lag,PV}} - t_{\mathrm{Lag,T}}} \frac{C_{\mathrm{A}}(\tau)}{1 - H_{\mathrm{LV}}}d\tau\right] \tag{11}$$

$$= \frac{(1-\gamma)K_{\mathrm{PV}}^{\mathrm{Trans}}}{1 - H_{\mathrm{LV}}}\left[\frac{v_{\mathrm{D}}\{1 - R_{\mathrm{PV}}(t)\}\{1 - R_{\mathrm{T}}(t)\}}{k_{\mathrm{PV}}}C_{\mathrm{A}}\left(t - t_{\mathrm{Lag,PV}} - t_{\mathrm{Lag,T}}\right) + \frac{F}{V_{\mathrm{T}}}R_{\mathrm{T,PV}}(t)I_{\mathrm{A}}\left(t - t_{\mathrm{Lag,PV}} - t_{\mathrm{Lag,T}}\right)\right],$$

where $R_{\mathrm{T,PV}}(t) = \frac{v_{\mathrm{D}}\{1-R_{\mathrm{T}}(t)\}R_{\mathrm{PV}}(t)}{F/V_{\mathrm{T}}} + \frac{\{1-R_{\mathrm{PV}}(t)\}R_{\mathrm{T}}(t)}{k_{\mathrm{PV}}}$. Consequently, a more stable estimation of the kinetic parameters is made possible even with low-temporal-resolution data because the solution allows estimating a continuously differentiable quantity. An example of the resultant fitting curves in HCC and normal tissue is shown in Fig. 2.

2.5 Image Preprocessing

Considering that CT images may suffer from low contrast-to-noise ratios as a consequence of the limitation of the radiation exposure, DCE-CT images were denoised by use of multiple observations Gaussian process regression with a 3×3 pixel kernel based on spatiotemporal information [16]. Next, to reduce movement-induced artifacts, we aligned the precontrast, arterial, and delayed-phase images (moving images) with the portal-phase image (fixed image). First, two images between the moving and fixed ones were registered by use of a 3D rigid transform, which, in turn, was used for initializing a registration with a 3D affine transform. The transform resulting from the affine registration was used as the bulk transform for a subsequent 3D deformable warping. Then, the symmetric force Demons deformable registration was performed by use of a multiresolution scheme [17]. Spatial filtering with a 5×5 median kernel was additionally applied to each dynamic series to compensate for a potential misalignment in pixel-wise correspondence between moving and fixed images, when we performed curve-fitting of the 4-phase DCE-CT data.

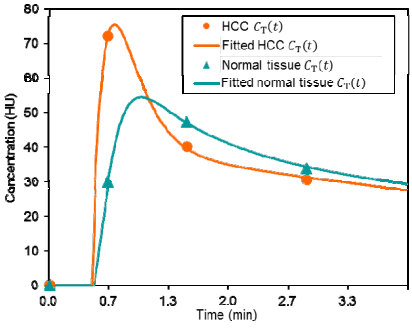


Fig. 2. Example of fitting HCC and normal tissue enhancement curves by the proposed liver tissue kinetic model in the 4-phase DCE-CT

3 Results

We investigated 8 patient HCC cases to illustrate the clinical applicability of the proposed continuous-time tracer kinetic model in 4-phase liver DCE-CT. The patients were scanned with a 64-row multidetector CT scanner (LightSpeed VCT or Discovery CT750 HD; GE Medical Systems, Milwaukee, WI). A total of 1.7 ml/kg (80 to 135 ml) of nonionic iodinated CA (Iomeron; Eisai, Tokyo, 350 mg/ml) was injected with 30 s injection duration at the rate of 3-5 ml/s and a volume as per 550-600 mgI/kg weight. The arterial-phase timing was determined with bolus tracking technology (Smart Prep; GE Healthcare), and the scan was initiated 17 s after the preselected threshold of 200 HU was attained, with a ROI placed in the aorta above the celiac axis branching, where the time lag of the bolus arrival to the aorta ($t_{\mathrm{Lag,A}}$) was determined by observation of a snapshot to show the onset time of temporal enhancement in the aorta. The portal venous phase was initiated 70 s, and the delayed phase 150 s after the preselected threshold of 200 HU was attained. The following CT parameters were used for obtaining volume data: 120 kVp, Auto mA, 16×0.625 mm detector collimation, 2.5 mm slice thickness, and a pitch of 1.

A typical HCC case is presented in Fig. 3, revealing that the kinetic parameter maps can be described adequately with the proposed model. The model parameters determined from ROI analysis in the HCC and normal tissue are summarized in Table 1. The ROIs were drawn manually over the primary HCCs and their adjacent normal tissues for each patient by an experienced radiologist. The ROI analysis indicated a tendency of considerably higher values for the total hepatic perfusion (F/V_{T}), arterial perfusion ($F_{\mathrm{A}}/V_{\mathrm{T}}$), and arterial flow fraction (γ) in the HCC than in the normal tissue, whereas the portal perfusion ($F_{\mathrm{PV}}/V_{\mathrm{T}}$), relative distribution volume (v_{D}) and time lag of bolus arrival from the input blood to the liver tissue ($t_{\mathrm{Lag,T}}$) did not show any obvious differences between the HCC and normal tissue. The increase in the arterial contribution to the HCC perfusion led to the increase in the arterial flow fraction in the HCC. This increased arterial perfusion coincides with the findings reported in previous studies of HCC with DCE-CT [18,19]. Also, despite the increase in the arterial flow fraction in HCC, there was still a tendency of slightly higher portal perfusion in the HCC than in normal tissue. This observation is in harmony with the finding that portal perfusion has been shown to decrease progressively with increasing

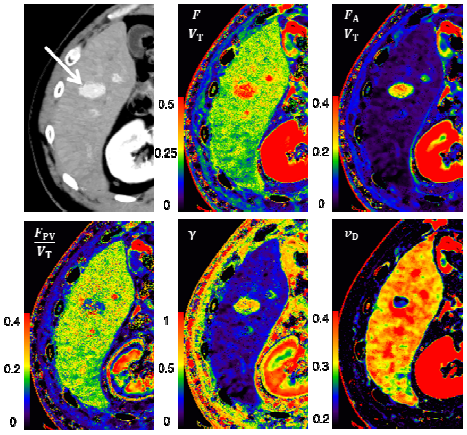


Fig. 3. Example of kinetic parameter maps by the proposed model in the 4-phase DCE-CT. The arrow indicates a HCC in the liver.

dedifferentiation of regenerating, dysplastic, and HCC nodules [20]. The normal tissue showed a tendency for slightly higher portal perfusion than the arterial perfusion, and thereby a lower arterial flow fraction than did the HCC, which is reasonable in the sense that about 75% of the incoming blood comes from the PV and the rest, 25%, results from the hepatic artery under normal conditions [21].

Table 1. Statistics (mean±SD) from ROI analysis of each kinetic parameter in the HCC and normal tissue of each study case

Patient no.		F/V_T (ml/min/ml)	F_A/V_T (ml/min/ml)	F_{PV}/V_T (ml/min/ml)	γ	v_D	$t_{Lag,T}$ (min)
1	HCC	0.24±0.26	0.23±0.17	0.02±0.1	0.98±0.09	0.25±0.03	0.17±0.1
	Normal	0.11±0.04	0.09±0.04	0.02±0.04	0.85±0.25	0.41±0.21	0.16±0.15
2	HCC	2.94±2.53	2.58±2.43	0.36±0.4	0.86±0.16	0.29±0.03	0.19±0.13
	Normal	0.31±0.16	0.17±0.11	0.14±0.12	0.57±0.27	0.27±0.02	0.13±0.12
3	HCC	0.52±0.87	0.25±0.18	0.27±0.69	0.56±0.15	0.35±0.09	0.13±0.09
	Normal	0.22±0.04	0.03±0.01	0.19±0.04	0.14±0.06	0.34±0.04	0.14±0.07
4	HCC	0.7±0.32	0.45±0.2	0.25±0.26	0.67±0.26	0.64±0.23	0.14±0.14
	Normal	0.3±0.64	0.06±0.03	0.24±0.08	0.2±0.12	0.36±0.03	0.11±0.09
5	HCC	0.52±0.71	0.38±0.51	0.14±0.27	0.79±0.24	0.21±0.04	0.07±0.15
	Normal	0.36±0.15	0.08±0.07	0.28±0.12	0.22±0.14	0.26±0.05	0.16±0.13
6	HCC	0.96±1.46	0.73±1.39	0.23±0.34	0.64±0.29	0.23±0.06	0.14±0.15
	Normal	0.31±0.2	0.09±0.1	0.22±0.14	0.3±0.22	0.26±0.02	0.09±0.08
7	HCC	2.52±3.06	2.48±3.08	0.04±0.07	0.94±0.09	0.31±0.03	0.11±0.08
	Normal	0.27±0.06	0.17±0.06	0.1±0.05	0.65±0.17	0.25±0.02	0.14±0.05
8	HCC	1.54±2.48	0.91±1.81	0.63±1.43	0.55±0.31	0.24±0.04	0.16±0.15
	Normal	0.64±0.38	0.3±0.42	0.33±0.2	0.41±0.3	0.29±0.02	0.35±0.08
HCC		1.24±1	1±0.97	0.24±0.19	0.75±0.17	0.32±0.14	0.14±0.04
Normal		0.31±0.15	0.12±0.09	0.19±0.1	0.42±0.25	0.3±0.06	0.16±0.08

4 Conclusion

We developed a novel computational solution for continuous-time modeling of tracer kinetics by using the convolution area property and the differentiation product rule without any discretization of the enhancement curves, thus enabling a more reliable estimation of the time lag between input and response enhancements as well as other kinetic parameters. The preliminary results of this study suggested that the proposed analytic scheme is feasible for estimation of kinetic parameters even in 4-phase liver DCE-CT, potentially being a practical guide for a tracer kinetic model-based curve-fitting, especially when data are temporally sparse.

References

1. Lee, T.Y., Purdie, T.G., Stewart, E.: CT imaging of angiogenesis. Q. J. Nucl. Med. 47(3), 171–187 (2003)
2. Lee, S.H., Cai, W., Yoshida, H.: Tracer kinetic modeling by morales-smith hypothesis in hepatic perfusion CT. In: Yoshida, H., Hawkes, D., Vannier, M.W. (eds.) Abdominal Imaging 2012. LNCS, vol. 7601, pp. 292–302. Springer, Heidelberg (2012)
3. Brix, G., Griebel, J., Kiessling, F., Wenz, F.: Tracer kinetic modelling of tumour angiogenesis based on dynamic contrast-enhanced CT and MRI measurements. Eur. J. Nucl. Med. Mol. Imaging 37(suppl. 1), S30–S51 (2010)
4. Konstas, A.A., Goldmakher, G.V., Lee, T.Y., Lev, M.H.: Theoretic basis and technical implementations of CT perfusion in acute ischemic stroke, part 1: Theoretic basis. AJNR Am. J. Neuroradiol. 30(4), 662–668 (2009)
5. Koh, T.S., Cheong, D.L., Hou, Z.: Issues of discontinuity in the impulse residue function for deconvolution analysis of dynamic contrast-enhanced MRI data. Magn. Reson. Med. 66(3), 886–892 (2011)
6. Lee, S.H., Kim, J.H., Kim, K.G., Park, S.J., Im, J.G.: Application of time sampling in brain CT perfusion imaging for dose reduction. In: Proc. SPIE Medical Imaging 2007: Physics of Medical Imaging, vol. 6510, pp. 65102P-1–65102P-10 (2007)
7. Bisdas, S., Foo, C.Z., Thng, C.H., Vogl, T.J., Koh, T.S.: Optimization of perfusion CT protocol for imaging of extracranial head and neck tumors. J. Digit. Imaging 22(5), 437–448 (2009)
8. Pandharipande, P.V., Krinsky, G.A., Rusinek, H., Lee, V.S.: Perfusion imaging of the liver: Current challenges and future goals. Radiology 234(3), 661–673 (2005)
9. Kambadakone, A.R., Sharma, A., Catalano, O.A., Hahn, P.F., Sahani, D.V.: Protocol modifications for CT perfusion (CTp) examinations of abdomen-pelvic tumors: Impact on radiation dose and data processing time. Eur. Radiol. 21(6), 1293–1300 (2011)
10. Catalano, O., Cusati, B., Sandomenico, F., Nunziata, A., Lobianco, R., Siani, A.: Multiple-phase spiral computerized tomography of small hepatocellular carcinoma: Technique optimization and diagnostic yield. Radiol. Med. 98(1-2), 53–64 (1999)
11. Kim, S.K., Lim, J.H., Lee, W.J., Kim, S.H., Choi, D., Lee, S.J., Lim, H.K., Kim, H.: Detection of hepatocellular carcinoma: Comparison of dynamic three-phase computed tomography images and four-phase computed tomography images using multidetector row helical computed tomography. J. Comput. Assist. Tomogr. 26(5), 691–698 (2002)
12. Orton, M.R., d'Arcy, J.A., Walker-Samuel, S., Hawkes, D.J., Atkinson, D., Collins, D.J., Leach, M.O.: Computationally efficient vascular input function models for quantitative kinetic modelling using DCE-MRI. Phys. Med. Biol. 53(5), 1225–1239 (2008)
13. Bae, K., Heiken, J., Brink, J.: Aortic and hepatic contrast medium enhancement at CT. I. Prediction with a computer model. Radiology 207, 647–655 (1998)
14. 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 diffusable tracer: Standardized quantities and symbols. J. Magn. Reson Imaging 10(3), 223–232 (1999)
15. Riad, S.M.: The deconvolution problem: An overview. Proc. IEEE 74(1), 82–85 (1986)
16. Zhu, F., Carpenter, T., Rodriguez Gonzalez, D., Atkinson, M., Wardlaw, J.: Computed tomography perfusion imaging denoising using gaussian process regression. Phys. Med. Biol. 57(12), N183–N198 (2012)

17. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: The ITK software guide. Kitware, Inc., Clifton Park (2005)
18. Miles, K.A., Hayball, M.P., Dixon, A.K.: Functional images of hepatic perfusion obtained with dynamic CT. Radiology 188(2), 405–411 (1993)
19. Tsushima, Y., Funabasama, S., Aoki, J., Sanada, S., Endo, K.: Quantitative perfusion map of malignant liver tumors, created from dynamic computed tomography data. Acad. Radiol. 11(2), 215–223 (2004)
20. Matsui, O., Kadoya, M., Kameyama, T., Yoshikawa, J., Takashima, T., Nakanuma, Y., Unoura, M., Kobayashi, K., Izumi, R., Ida, M.: Benign and malignant nodules in cirrhotic livers: Distinction based on blood supply. Radiology 178(2), 493–497 (1991)
21. Chiandussi, L., Greco, F., Sardi, G., Vaccarino, A., Ferraris, C.M., Curti, B.: Estimation of hepatic arterial and portal venous blood flow by direct catheterization of the vena porta through the umbilical cord in man. Preliminary results. Acta Hepatosplenol. 15(3), 166–171 (1968)

Use of Tracer Kinetic Model-Driven Biomarkers for Monitoring Antiangiogenic Therapy of Hepatocellular Carcinoma in First-Pass Perfusion CT

Sang Ho Lee[1], Koichi Hayano[2], Dushyant Sahani[2], and Hiroyuki Yoshida[1]

[1] 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
[2] Division of Abdominal Imaging and Intervention, Department of Radiology,
Massachusetts General Hospital, 55 Fruit St., Boston, Massachusetts 02114, USA
{khayano,dsahani}@partners.org

Abstract. Development of vascularly targeted anti-cancer therapies has led to an interest in determining the in vivo effectiveness of anti-tumor agents in patients. As the antiangiogenic agents may have significant effects without causing tumor shrinkage, their microcirculatory characteristics have the potential to be response biomarkers. Perfusion CT (PCT) studies can quantify the microcirculatory status of liver tumors, and can be used for assessing the effectiveness of antiangiogenic therapy. Our purpose in this study was to compare five different tracer kinetic models for the analysis of first-pass hepatic PCT data, to investigate whether kinetic parameters differ in significance among different kinetic models, and to select the best single prognostic biomarker with respect to the prediction of 6-month progression-free survival (PFS) of patients with advanced hepatocellular carcinoma (HCC). The first-pass PCT was performed at baseline and on days 10 to 12 after initiation of antiangiogenic treatment. The PCT data were analyzed retrospectively by the Tofts-Kety (TK), extended TK (ETK), two-compartment exchange (2CX), adiabatic approximation to tissue homogeneity (AATH), and distributed parameter (DP) models. Kinetic parameters consisted of blood flow (BF), blood volume (BV), mean transit time (MTT), and permeability-surface area product (PS), mean values of which within HCC were compared between baseline and post-treatment by the Wilcoxon signed-rank test. Baseline mean kinetic parameters within HCC and relative percent changes (%changes) in the mean and standard deviation (SD) from baseline to post-treatment were also compared in terms of PFS discrimination by use of Spearman correlation analysis. After treatment, the changes in the kinetic parameter values were significantly different among models. The results suggested that the %change of SD for BV is an effective prognosis biomarker, potentially reflecting that treatment-induced change of vascular heterogeneity plays a role in the assessment of the HCC response. Based on the predictive ranking of a single biomarker, the AATH model was the best predictor of 6-month PFS in the first-pass PCT analysis.

Keywords: Hepatocellular carcinoma, antiangiogenic treatment, image biomarker, perfusion CT.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 270–279, 2013.

1 Introduction

Hepatocellular carcinoma (HCC) is the third most common cause of cancer-related death worldwide and is responsible for more than 500,000 deaths each year [1]. Because many patients with early disease are asymptomatic [2], HCC is frequently diagnosed at an advanced stage, by which time it is often untreatable [3]. HCC is a hypervascular tumor type with increased levels of vascular endothelial growth factor and microvessel density, which derive neovasculature through the process of angiogenesis [4], [5].

Advances in molecular medicine offer the potential to move beyond traditional cytotoxic anticancer treatments. Treatment with molecular targeted agents such as antiangiogenic agents holds great promise for inhibiting the growth of highly vascularized tumors. As the antiangiogenic agents may have significant effects without causing tumor shrinkage, the microcirculatory characteristics have the potential to be response biomarkers. Tumor angiogenesis is a complex biological process that is critical for tumor growth and metastasis [6]. Therefore, quantifying tumor angiogenesis is important for risk stratification, evaluation of disease progression, and monitoring of the response to therapy [7]. Consequently, an accurate noninvasive method for addressing the complexities of tumor angiogenesis associated with the expression of suspected biomarkers would be highly desirable.

Perfusion CT (PCT) is increasingly being used for quantifying tumor vascularity and monitoring of the antiangiogenic response in HCC and various other solid tumors [8]. Because of the advantage of its rapid scanning speed as well as high spatial and temporal resolution, PCT imaging is gaining popularity for acquisition of quantitative measurements of various tissue kinetic parameters by modeling of tracer kinetics [9], [10]. The kinetic parameters can indirectly reflect the tumor vascularity as well as early antiangiogenic effects. To date, there has been an effort to seek data comparing the relative performance of each tracer kinetic model in tumors in terms of the prediction of clinical outcome. Several studies have evaluated how kinetic parameter values differ among some kinetic models, which have indicated that kinetic parameter values derived from different models are not necessarily similar or significantly different [11], [12]. So far, no studies have determined whether the choice of a kinetic model has the potential to affect the prediction of the clinical outcome as a prognostic biomarker in HCC. Thus, our aim was to compare five different tracer kinetic models for the analysis of first-pass PCT (i.e., dynamic CT scanning during the first circulation of an injected contrast agent (CA) through the body) data, to investigate whether kinetic parameters have different significance among different tracer kinetic models, and to select the best single prognostic biomarker, with respect to the prediction of 6-month progression-free survival (PFS) in advanced HCC.

2 Methods

2.1 Single-Input Two Compartment Model

The liver tissue is essentially supplied by both the hepatic artery and the portal vein (PV). However, one of the main limitations of PCT is that it would be obtained with

only a limited number of slices, i.e., only a small volume. In addition, HCC with a PV tumor thrombus can directly invade or extrinsically compress the PV, often making the PV invisible on PCT images. Thus, in this study, we considered the CA-blood concentration-time curve only in the hepatic artery, $C_{\mathrm{A}}(t)$ (in g/ml), as a single input for liver tissue enhancement during the early arterial phase (first-pass), by assuming that the PV contribution to the early liver tissue enhancement is negligible as the PV enhancement can be delayed in its onset compared to the arterial enhancement. The concentration of CA in liver tissue, $C_{\mathrm{T}}(t)$, during the first pass can be described by the following expression:

$$C_{\mathrm{T}}(t) = \frac{F}{V_{\mathrm{T}}} R_{\mathrm{T}}(t) \otimes \frac{C_{\mathrm{A}}(t)}{1 - H_{\mathrm{LV}}}, \tag{1}$$

where $R_{\mathrm{T}}(t)$, F (in ml/min), V_{T} (in ml), H_{LV} and $\otimes$ denote the impulse tissue residue response function, hepatic plasma flow, tissue volume, hematocrit of blood in large vessels ($\cong 0.45$) [13], and convolution operator, respectively. Thus, F/V_{T} (ml/min/ml) becomes the hepatic perfusion. In this study, we compared five different tracer kinetic models of the Tofts-Kety (TK), extended TK (ETK), two compartment exchange (2CX), adiabatic approximation to the tissue homogeneity (AATH), and distributed parameter (DP) models for $R_{\mathrm{T}}(t)$ [14]. Each model represented its own $R_{\mathrm{T}}(t)$ in two compartments: plasma space and interstitial space. The impulse tissue residue functions $R_{\mathrm{T}}(t)$ for the TK, ETK, 2CX, AATH, and DP models are given by the following expressions, respectively:

$$R_{\mathrm{T,TK}}(t) = Ee^{-\frac{v_{\mathrm{P}}EF}{v_{\mathrm{I}}V_{\mathrm{P}}}t}, \tag{2}$$

$$R_{\mathrm{T,ETK}}(t) = \frac{V_{\mathrm{P}}}{F}\delta(t) + Ee^{-\frac{v_{\mathrm{P}}EF}{v_{\mathrm{I}}V_{\mathrm{P}}}t}, \tag{3}$$

$$R_{\mathrm{T,2CX}}(t) = Ae^{\alpha t} + (1-A)e^{\beta t} \text{ with } \begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \frac{1}{2}\left[-\left\{\frac{F}{V_{\mathrm{P}}} + \left(1+\frac{v_{\mathrm{P}}}{v_{\mathrm{I}}}\right)\frac{PS}{V_{\mathrm{P}}}\right\} \pm \sqrt{\left\{\frac{F}{V_{\mathrm{P}}} + \left(1+\frac{v_{\mathrm{P}}}{v_{\mathrm{I}}}\right)\frac{PS}{V_{\mathrm{P}}}\right\}^2 - 4\frac{v_{\mathrm{P}}}{v_{\mathrm{I}}}\frac{F}{V_{\mathrm{P}}}\frac{PS}{V_{\mathrm{P}}}}\right] \text{ and } A = \frac{\alpha + \left(1+\frac{v_{\mathrm{P}}}{v_{\mathrm{I}}}\right)\frac{PS}{V_{\mathrm{P}}}}{\alpha - \beta}, \tag{4}$$

$$R_{\mathrm{T,AATH}}(t) = u(t) + \left[Ee^{-\frac{v_{\mathrm{P}}EF}{v_{\mathrm{I}}V_{\mathrm{P}}}\left(t-\frac{V_{\mathrm{P}}}{F}\right)} - 1\right]u\left(t - \frac{V_{\mathrm{P}}}{F}\right), \tag{5}$$

$$R_{\mathrm{T,DP}}(t) = u(t) + e^{-\frac{PS}{F}}u\left(t-\frac{V_{\mathrm{P}}}{F}\right) \cdot \int_0^{t-\frac{V_{\mathrm{P}}}{F}} e^{-\frac{v_{\mathrm{P}}PS}{v_{\mathrm{I}}V_{\mathrm{P}}}\tau}\left[\delta(\tau) + \frac{PS}{V_{\mathrm{P}}}\sqrt{\frac{v_{\mathrm{P}}}{v_{\mathrm{I}}}\frac{V_{\mathrm{P}}}{F}\frac{1}{\tau}}I_1\left(2\frac{PS}{V_{\mathrm{P}}}\sqrt{\frac{v_{\mathrm{P}}}{v_{\mathrm{I}}}\frac{V_{\mathrm{P}}}{F}\tau}\right)\right]d\tau, \tag{6}$$

where $R_{\mathrm{T,TK}}(t)$, $R_{\mathrm{T,ETK}}(t)$, $R_{\mathrm{T,2CX}}(t)$, $R_{\mathrm{T,AATH}}(t)$ and $R_{\mathrm{T,DP}}(t)$ represent the impulse tissue residue functions of the TK, ETK, 2CX, AATH, and DP models, respectively. Note that $\delta(t)$, $u(t)$, $I_1(t)$, $V_{\mathrm{P}} = v_{\mathrm{P}}V_{\mathrm{T}}$, v_{P}, v_{I}, and $E = 1 - e^{-PS/F}$,

where PS (in ml/min) is the permeability surface area product, denote the Dirac delta function, unit step function, modified Bessel function of the first kind, plasma volume, plasma volume fraction, interstitial volume fraction, and extraction fraction, respectively. As $C_{\mathrm{A}}(t)$ is sampled at a major feeding artery (abdominal aorta) upstream of the liver tissue, the bolus arrival time in $C_{\mathrm{A}}(t)$ is earlier than that of $C_{\mathrm{T}}(t)$. To account for the difference in bolus arrival times, a time lag (delay) to the liver tissue, $t_{\mathrm{Lag,T}}$ can be imposed on $R_{\mathrm{T}}(t)$ in the calculation of $C_{\mathrm{T}}(t)$, i.e.: $C_{\mathrm{T}}(t) = \frac{F}{V_{\mathrm{T}}} R_{\mathrm{T}}\left(t - t_{\mathrm{Lag,T}}\right) \otimes \frac{C_{\mathrm{A}}(t)}{1-H_{\mathrm{LV}}}$. The convolution operation can be visualized as a reflection of $R_{\mathrm{T}}(t)$ about $t = 0$ and by summing of the overlapping areas of the reflected $R_{\mathrm{T}}(t)$ and $C_{\mathrm{A}}(t)$ as the reflected $R_{\mathrm{T}}(t)$ is shifted along the positive t direction. The process of determining $R_{\mathrm{T}}(t)$, given $C_{\mathrm{A}}(t)$ and $C_{\mathrm{T}}(t)$, is called deconvolution. However, it is clear that deconvolution cannot give a unique answer because there would certainly be more than one $R_{\mathrm{T}}(t)$ that would give a good approximation to $C_{\mathrm{T}}(t)$ after convolution with $C_{\mathrm{A}}(t)$. Thus, considering computational simplicity and that the temporal resolution of the PCT data analyzed was high enough (≅0.6 s), we performed a trapezoidal convolution operation to approximate $C_{\mathrm{T}}(t)$. For a PCT dataset consisting of N images obtained with a time interval $\Delta t_k = t_{k+1} - t_k$ for k =0, ⋯, $N-1$, the trapezoidal convolution approximation is given in the following form:

$$C_{\mathrm{T}}(t_k) \cong \frac{F}{V_{\mathrm{T}}} \sum_{i=0}^{k-1} \Big[R_{\mathrm{T}}\left(t_i + t_{\mathrm{Lag,T}}\right) C_{\mathrm{A}}(t_k - t_i) + R_{\mathrm{T}}\left(t_{i+1} + t_{\mathrm{Lag,T}}\right) C_{\mathrm{A}}(t_k - t_{i+1}) \Big] \frac{\Delta t_i}{2(1 - H_{\mathrm{LV}})}. \tag{7}$$

Note that $R_{\mathrm{T}}(t_k) = 0$ for $t_k \le t_{\mathrm{Lag,T}}$, and $t_0 = 0$ is the first precontrast phase of the PCT. With this approach, an estimate for the following set of kinetic parameters was obtained from the enhancement curves of $C_{\mathrm{T}}(t)$ and $C_{\mathrm{A}}(t)$: blood flow (BF), blood volume (BV), mean transit time (MTT), and permeability-surface area product (PS). The BV (in ml/100 g), BF (in ml/min/100 g), MTT (min), and PS (in ml/min/100 g) were computed according to $BV = \frac{V_{\mathrm{P}}}{(1-H_{\mathrm{SV}})m} \cdot 100 = \frac{v_{\mathrm{P}}}{(1-H_{\mathrm{SV}})\rho_{\mathrm{T}}} \cdot 100$, $BF = BV \frac{F}{V_{\mathrm{P}}}$, $PS = BV \frac{PS}{V_{\mathrm{P}}}$, and $MTT = \frac{V_{\mathrm{P}} + V_{\mathrm{I}}}{F}$, where H_{SV} is the hematocrit in small vessels (≅0.25), $m = \rho_{\mathrm{T}} V_{\mathrm{T}}$ is the mass of the tissue with density ρ_{T} (=1.04 g/cm^3 in the case of soft tissues), and $V_{\mathrm{I}} = v_{\mathrm{I}} V_{\mathrm{T}}$ the interstitial volume.

2.2 Arterial Input Function Model

As a general bolus model, a gamma-variate (GV) function was fitted to the first-pass portion of $C_{\mathrm{A}}(t)$, given by the following expression:

$$C_{\mathrm{A}}(t) = a_{\mathrm{B}}\left(t - t_{\mathrm{Lag,A}}\right)^{\kappa} e^{-\mu_{\mathrm{B}}\left(t - t_{\mathrm{Lag,A}}\right)}, \tag{8}$$

where a_{B} is a scaling factor, κ and μ_{B} determine the bolus shape, and $t_{\mathrm{Lag,A}}$ is the arrival time of the bolus to the artery. Note that $C_{\mathrm{A}}(t) = 0$ for $t \leq t_{\mathrm{Lag,A}}$. The last fitting point for the GV was determined from the time point at half the bolus depth after the peak of the arterial enhancement [15].

2.3 Image Preprocessing

Considering that PCT images suffer from low contrast-to-noise ratios as a consequence of the limitation of the patient's exposure to radiation, the PCT images were denoised by use of multiple observations Gaussian process regression with a 3×3 pixel kernel based on spatiotemporal information [16]. Next, to reduce movement-induced artifacts, we aligned each set of PCT images (moving images) to the first image as a template (fixed image). First, two images between the moving and fixed ones were roughly aligned by a transform initialization based on image intensity moments; then they were registered by use of a 3D rigid transform, that, in turn, was used for initializing of a registration with a 3D affine transform. The transform resulting from the affine registration was used as the bulk transform for a subsequent 3D deformable transform. B-spline deformable registration was computed, and finally the resulting transform was used for resampling of the moving images [17]. Because the PCT data obtained had high temporal resolution, the kinetic parameters were calculated based on a pixel-by-pixel analysis. To draw a region of interest (ROI) for HCC, an experienced surgeon drew an ROI on each kinetic parameter map, and corresponding parameter values were averaged for all ROIs in each patient.

2.4 Statistical Analysis

A paired comparison of the mean values of the kinetic parameters (i.e., BF, BV, MTT, and PS) within HCC between baseline (pre-treatment) and post-treatment was evaluated for each tracer kinetic model by use of the Wilcoxon signed-rank test in our effort to investigate whether early antiangiogenic changes were significant after treatment regardless of the different PFS groups. One-way analysis of variance (Friedman test) was performed for comparison of the mean values of the kinetic parameters among the five different tracer kinetic models (i.e., TK, ETK, 2CX, AATH, and DP models) at each baseline and post-treatment. The baseline mean values of the kinetic parameters within HCC and relative percent changes (%changes) in the mean and standard deviation (SD) values from baseline to post-treatment were compared with respect to PFS discrimination by use of Spearman correlation analysis. The %changes in the mean and SD values of the kinetic parameters were given as $\left(mean_{\mathrm{post}} - mean_{\mathrm{pre}}\right) \cdot 100/mean_{\mathrm{pre}}$ and $\left(SD_{\mathrm{post}} - SD_{\mathrm{pre}}\right) \cdot 100/SD_{\mathrm{pre}}$, where $mean_{\mathrm{pre}}$, SD_{pre}, $mean_{\mathrm{post}}$, and SD_{post} represent the mean and SD values of the kinetic parameters at baseline and post-treatment, respectively. P <0.05 was considered statistically significant.

3 Results

A total of 19 patients with advanced HCC were included in this study. The PFS rate at 6 months of study cohort was 53%. Therefore, patients were separated into 2 groups: patients with PFS >6 months (n=10) and those with PFS ≤6 months (n=9) for this analysis. The patients were treated with bevacizumab at a dose of 10 mg/kg i.v. on day 1 of cycle 1 (14 days). For the subsequent 28-day cycle, patients were treated with bevacizumab at 10 mg/kg on days 1 and 15, gemcitabine at 1000 mg/m^2 i.v. at a dose rate infusion of 10 mg/m^2/min on days 2 and 16, and oxaliplatin at 85 mg/m^2 at a 2-hour i.v. infusion on days 2 and 16 of every cycle (GEMOX-B chemotherapy). The dose of bevacizumab was fixed at 10 mg/kg. The PCT was performed at baseline and on days 10 to 12 after initiation of antiangiogenic treatment. For PCT imaging, a 16-section multidetector CT scanner (LightSpeed; GE Medical Systems, Milwaukee, WI) was used. For initial localization of the HCC, an unenhanced CT scan of the abdomen was obtained during an end-expiratory breath-hold, and then a 2-cm tumor region was selected independently for PCT imaging in the maximal diameter of the HCC. Four consecutive slices in the region of the known HCC were selected for cine image acquisition. A dynamic contrast-enhanced study of the selected four slices was performed in a single breath-hold of 30 to 35 s at the end of expiration with a static table position. The duration of 30 to 35 s for PCT scanning was chosen on the basis of patient maximum breath-hold capacity and to minimize the radiation dose and respiratory artifacts. A total of 50 to 70 ml of nonionic iodinated CA (Isovue; Bracco, Princeton, NJ) (300 mgI/ml) was injected at a rate of 5 to 7 ml/s through an 18-gauge intravenous cannula. The following CT parameters were used for acquisition of PCT data: rotation time, 0.5 s; cine acquisition at every 0.6 s; 100 to 120 kVp; 200 to 240 mA; 4 sections per gantry rotation; and 5-mm reconstructed section thickness. Scanning was initiated after a 5 s delay from the start of injection, and images were obtained for a total duration of 30 to 35 s.

Fig. 1 shows an example of the kinetic parameter maps (i.e., BF, BV, MTT, and PS) derived from the five different kinetic models (i.e., TK, ETK, 2CX, AATH, and DP) before and 2 weeks after antiangiogenic treatment in a 73-year-old woman with PFS >6 months. The kinetic parameter changes were substantial, with a decrease in tumor BF, BV, and PS, and an increase in MTT from their baseline values. This trend was similar over the five different kinetic models even though the absolute values of the kinetic parameters were different among the different models. However, it should be noted that this case was a progressive disease after treatment in terms of tumor burden measurement as defined by RECIST (Response Evaluation Criteria in Solid Tumors) [18]. This mismatch between the kinetic parameter estimates and the clinical response indicates that the kinetic parameters would be more sensitive image biomarkers for monitoring of early antiangiogenic treatment effects as well as in predicting PFS, as compared with RECIST [10].

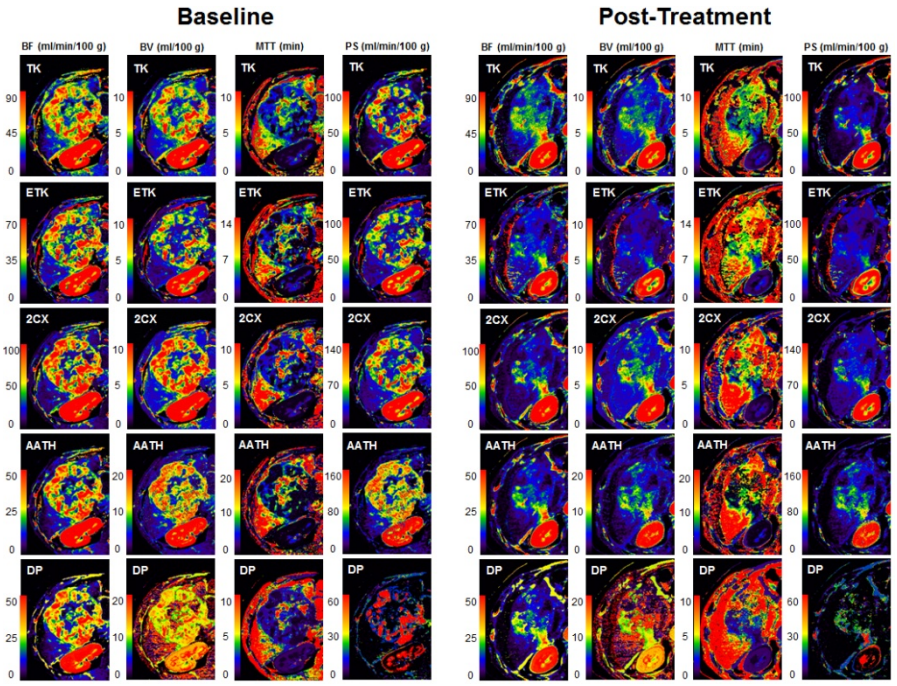


Fig. 1. Example of kinetic parameter maps (BF, BV, MTT, and PS) derived from five different tracer kinetic models (TK, ETK, 2CX, AATH, and DP) before and 2 weeks after antiangiogenic treatment in a 73-year-old woman with PFS >6 months. At baseline, the HCC in the right lobe of the liver shows increased vascularity, shown as a distinct range of colors compared with background liver tissue. Following therapy, there is a dramatic reduction in the BF, BV, and PS, and an increase in the MTT on the colored maps.

The mean values of the kinetic parameters within HCC for the five different kinetic models are summarized in Table 1, being compared between baseline and post-treatment regardless of the two different PFS groups. The results of the Friedman test revealed that all kinetic parameters analyzed were statistically significantly different ($P<0.001$) among the five different models of both baseline and post-treatment. The BF was higher with the 2CX model than with the rest of the models at both baseline and post-treatment, whereas it was lower with the DP model. The BV was higher with the AATH model than with the rest of the models at baseline, and higher with the DP model at post-treatment, whereas it was lower with the ETK model at both baseline and post-treatment. The MTT was shorter with the TK model than with the rest of the models at both baseline and post-treatment, whereas it was longer with the AATH model. The PS was higher with the AATH model than with the rest of the models at both baseline and post-treatment, whereas it was lower with the DP model. The results of the Wilcoxon signed-rank test revealed whether the changes of the kinetic parameters were statistically significant after antiangiogenic treatment. Between baseline and post-treatment, the BF, BV, PS, and MTT were statistically significantly different in the TK, 2CX, and AATH models, and the BF, PS, and MTT except the

BV in the DP model. However, the ETK model did not show statistical significance for BF, BV, and PS. The MTT was statistically significantly different in all kinetic models considered. After treatment, the kinetic parameter changes in all kinetic models except the ETK model were substantial, with decreases in the BF, BV, and PS, and an increase in the MTT from their baseline values like other previous studies [10], [19], [20].

Table 1. Early response: kinetic parameters at baseline and following antiangiogenic treatment at 2 week (n=19)

Kinetic parameter	Kinetic model	Baseline (mean±SD)	Post-treatment (mean±SD)	P-value†	P-value‡	
					Baseline	Post-treatment
BF (ml/min/100 g)	TK	55.38±16.13	47.27±19.36	0.004*	<0.001*	<0.001*
	ETK	42.9±13.31	44.64±28.22	0.l98		
	2CX	62.49±26.04	49.41±28.83	0.007*		
	AATH	27.95±10.17	22.29±12.46	0.002*		
	DP	26.22±8.65	21.04±9.9	0.015*		
BV (ml/100 g)	TK	5.6±1.6	4.85±1.89	0.004*	<0.001*	<0.001*
	ETK	4.61±1.39	4.77±2.93	0.156		
	2CX	7.05±2.65	5.83±3.36	0.005*		
	AATH	11.84±4.14	9.47±5.19	0.003*		
	DP	11.82±1.48	11.74±1.74	0.311		
MTT (min)	TK	3.09±1.03	3.9±1.48	0.015*	<0.001*	<0.001*
	ETK	4.09±1.54	5.1±2.04	0.003*		
	2CX	3.32±1.67	4.39±2.05	0.015*		
	AATH	6.14±3.29	7.79±3.62	0.025*		
	DP	3.29±1.98	5.52±4.38	0.006*		
PS (ml/min/100 g)	TK	47.23±18.92	39.05±22	0.008*	<0.001*	<0.001*
	ETK	43.53±13.84	44.59±28.61	0.156		
	2CX	69.05±26.88	56.12±34.03	0.005*		
	AATH	87.79±31.39	69.26±39.44	0.003*		
	DP	20.82±17.9	13.24±12.61	0.004*		

Asterisk (*): statistical significance (P<0.05), †Wilcoxon signed-rank test, ‡Friedman test.

The results of the Spearman correlation analysis with respect to PFS discrimination are summarized in Table 2. The baseline mean values of the kinetic parameters were not statistically significant in any of the kinetic models. The %changes of the mean for the BF and PS were statistically significantly correlated with a longer PFS only in the ETK model. The %changes in the SD for the BF (ρ=0.404), BV (ρ=0.52), and PS (ρ=0.481) were statistically significantly correlated with a longer PFS in the AATH model, those for the BF (ρ=0.404) and BV (ρ=0.404) in the ETK model, and those for the BV (ρ=0.423) and PS (ρ=0.404) in the 2CX model. Overall, the effective prognostic biomarkers with P<0.05 were in the range of $0.404 \leq \rho \leq 0.52$ in the

Spearman correlation coefficients. The %change in the SD presented a relatively better correlation in terms of PFS discrimination.

Table 2. Spearman correlation coefficients (ρ) of kinetic parameters in HCC with respect to PFS (n=19) discrimination

Kinetic parameter	Kinetic model	Baseline mean	%change of mean	%change of SD
BF (ml/min/100 g)	TK	0.173	0.212	0.269
	ETK	0.231	0.404*	0.404*
	2CX	0.192	0.173	0.289
	AATH	0.173	0.25	0.404*
	DP	0.269	0.192	0.135
BV (ml/100 g)	TK	0.212	0.115	0.25
	ETK	0.192	0.385	0.404*
	2CX	0.154	0.154	0.423*
	AATH	0.115	0.327	0.52*
	DP	0.269	-0.327	-0.192
MTT (min)	TK	-0.173	-0.154	0.058
	ETK	-0.135	-0.077	0.077
	2CX	-0.154	-0.077	0.058
	AATH	-0.173	0.231	0.058
	DP	-0.096	-0.077	0.135
PS (ml/min/100 g)	TK	0.192	0.115	0.231
	ETK	0.25	0.404*	0.327
	2CX	0.154	0.212	0.404*
	AATH	0.115	0.308	0.481*
	DP	0.25	-0.154	0

Asterisk (*): statistical significance (one-sided P<0.05)

4 Conclusions

The preliminary results suggest that the kinetic parameter values were significantly different between the different tracer kinetic models, and that the %change in the SD for BV is an effective prognosis biomarker, potentially reflecting that a treatment-induced change in vascular heterogeneity plays an important role in the assessment of the HCC response. Based on the predictive ranking of a single biomarker, the AATH model was the best predictor of 6-month PFS in the first-pass PCT analysis.

References

1. Jemal, A., Siegel, R., Xu, J., Ward, E.: Cancer Statistics, 2010. CA Cancer J. Clin. 60, 277–300 (2010)
2. Kew, M.C., Dos Santos, H.A., Sherlock, S.: Diagnosis of Primary Cancer of the Liver. Br. Med. J. 4, 408–411 (1971)

3. A New Prognostic System for Hepatocellular Carcinoma: A Retrospective Study of 435 Patients: The Cancer of the Liver Italian Program (CLIP) Investigators. Hepatology 28, 751–755 (1998)
4. Yamaguchi, R., Yano, H., Iemura, A., Ogasawara, S., Haramaki, M., Kojiro, M.: Expression of Vascular Endothelial Growth Factor in Human Hepatocellular Carcinoma. Hepatology 28, 68–77 (1998)
5. Messerini, L., Novelli, L., Comin, C.E.: Microvessel Density and Clinicopathological Characteristics in Hepatitis C Virus and Hepatitis B Virus Related Hepatocellular Carcinoma. J. Clin. Pathol. 57, 867–871 (2004)
6. Lee, T.Y., Purdie, T.G., Stewart, E.: CT Imaging of Angiogenesis. Q. J. Nucl. Med. 47, 171–187 (2003)
7. Miles, K.A.: Functional Computed Tomography in Oncology. Eur. J. Cancer 38, 2079–2084 (2002)
8. Kan, Z., Kobayashi, S., Phongkitkarun, S., Charnsangavej, C.: Functional CT Quantification of Tumor Perfusion After Transhepatic Arterial Embolization in a Rat Model. Radiology 237, 144–150 (2005)
9. Kambadakone, A.R., Sahani, D.V.: Body Perfusion CT: Technique, Clinical Applications, and Advances. Radiol. Clin. North Am. 47, 161–178 (2009)
10. Jiang, T., Kambadakone, A., Kulkarni, N.M., Zhu, A.X., Sahani, D.V.: Monitoring Response to Antiangiogenic Treatment and Predicting Outcomes in Advanced Hepatocellular Carcinoma using Image Biomarkers, CT Perfusion, Tumor Density, and Tumor Size (RECIST). Invest. Radiol. 47, 11–17 (2012)
11. Goh, V., Halligan, S., Bartram, C.I.: Quantitative Tumor Perfusion Assessment with Multidetector CT: Are Measurements from Two Commercial Software Packages Interchangeable? Radiology 242, 777–782 (2007)
12. Koh, T.S., Ng, Q.S., Thng, C.H., Kwek, J.W., Kozarski, R., Goh, V.: Primary Colorectal Cancer: Use of Kinetic Modeling of Dynamic Contrast-Enhanced CT Data to Predict Clinical Outcome. Radiology 267, 145–154 (2013)
13. Brix, G., Griebel, J., Kiessling, F., Wenz, F.: Tracer Kinetic Modelling of Tumour Angiogenesis Based on Dynamic Contrast-Enhanced CT and MRI Measurements. Eur. J. Nucl. Med. Mol. Imaging 37(suppl. 1), S30–S51 (2010)
14. Lee, S.H., Cai, W., Yoshida, H.: Tracer Kinetic Modeling by Morales-Smith Hypothesis in Hepatic Perfusion CT. In: Yoshida, H., Hawkes, D., Vannier, M.W. (eds.) Abdominal Imaging 2012. LNCS, vol. 7601, pp. 292–302. Springer, Heidelberg (2012)
15. Thompson Jr., H.K., Starmer, C.F., Whalen, R.E., McIntosh, H.D.: Indicator Transit Time Considered as a Gamma Variate. Circ. Res. 14, 502–515 (1964)
16. Zhu, F., Carpenter, T., Rodriguez Gonzalez, D., Atkinson, M., Wardlaw, J.: Computed Tomography Perfusion Imaging Denoising using Gaussian Process Regression. Phys. Med. Biol. 57, N183–N198 (2012)
17. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: The ITK Software Guide. Kitware, Inc., Clifton Park (2005)
18. Eisenhauer, E.A., Therasse, P., Bogaerts, J., Schwartz, L.H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., Verweij, J.: New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (Version 1.1). Eur. J. Cancer 45, 228–247 (2009)
19. Sahani, D.V., Kalva, S.P., Hamberg, L.M., Hahn, P.F., Willett, C.G., Saini, S., Mueller, P.R., Lee, T.Y.: Assessing Tumor Perfusion and Treatment Response in Rectal Cancer with Multisection CT: Initial Observations. Radiology 234, 785–792 (2005)
20. Zhang, Q., Yuan, Z.G., Wang, D.Q., Yan, Z.H., Tang, J., Liu, Z.Q.: Perfusion CT Findings in Liver of Patients with Tumor during Chemotherapy. World J. Gastroenterol. 16, 3202–3205 (2010)

A Statistical Shape Model for Multiple Organs Based on Synthesized-Based Learning

Atsushi Saito[1], Misaki Nakada[1], Elco Oost[1],
Akinobu Shimizu[1], Hidefumi Watanabe[1], and Shigeru Nawano[2]

[1] Tokyo University of Agriculture and Technology, Koganei, Tokyo, Japan
[2] Center for Radiological Science, Mita Hopital,
International University of Health and Welfare, Minato-ku, Tokyo, Japan
{50013834203,simiz,wtnbhdf}@st.tuat.ac.jp,
c_r_oost@hotmail.com, snawano@iuhw.ac.jp

Abstract. This paper presents a statistical shape model for multiple abdominal organs using synthesized-based learning to compensate the lack of a large manually labeled training data set. Experiments on 23 non-contrast CT volumes showed that a model trained on both true and synthesized data, outperforms conventional shape models, in terms of generalization, specificity and overlap of neighboring organs.

Keywords: Synthesized-based learning, level set morphing, statistical shape model, multiple abdominal organs, CT volume.

1 Introduction

Multi-organ segmentation of abdominal CT volumes is a growing topic in the field of computer aided diagnosis and computer-assisted intervention [1-6]. In the last decade, statistical shape models (SSMs) have become a powerful tool for automatic segmentation of organs. Training of an SSM requires a substantial set of images and accompanying manually drawn label images of the organs of interest. However, manual labeling of medical image data is time consuming and thus it is difficult to generate a sufficiently large training data set. Consequently, lack of sufficient training data causes deterioration in the performance of the SSM. This is particularly true for multi-organ SSMs, which require manual labeling of multiple organs. This paper investigates how to boost the training of an SSM, when only limited training data is available.

Inspired by synthesized-based learning [7], this paper presents a novel algorithm that generates a large number of additional, synthetic training data by interpolating between two training samples of multiple abdominal organs. The proposed interpolation is based on level set morphing [8] that was originally developed for interpolation between two resembling objects. We extend the method to handle multiple organs, and incorporate a scheme to eliminate possible overlap between adjacent organs. To the best of our knowledge, this is the first paper in medical image processing that proposes synthesized-based learning to generate additional training data for an SSM.

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 280–289, 2013.

A number of papers have reported on multiple organ SSMs. Duta et al. [9] presented a point distribution model (PDM) for multiple anatomical brain structures and Okada et al. [3] created a PDM based multi organ model for seven abdominal organs. The level set based SSM, or signed distance based SSM, is another promising method and has the benefit that it does not require point correspondence between the training samples. Tsai et al. developed an implicit multi-shape parametric model [10] for three brain structures, by concatenating three vectorized level set functions and applying principal component analysis (PCA) to the set of the concatenated vectors of training subjects. A limitation of this approach is that it requires a large number of manually labeled training samples. Using only a limited number of training samples, both the shape and pose description of the organs will be poor, increasing also the risk of overlapping organs in the model. Several multi-organ modeling methods exist that rule out organ overlap, such as e.g. the LogOdds based method by Pohl *et al.* [11], the label space approach by Malcom *et al.* [12] and the ILR transform for compositional data, introduced by Changizi *et al.* [13]. All these methods compute, for all voxels and all organs, the probability that a voxel represents a specific organ. Then, every voxel is assigned the label of the organ with the highest probability. This results in a one-label-per-voxel description, intrinsically avoiding organ overlap. However, mainly on the boundaries of touching organs, this can lead to unnatural shape representations.

The work presented in this paper focuses on a level set based SSM for modeling the abdominal organs, because of the simplicity in constructing the model and because of the accuracy in the shape description. To overcome the limitations of scarce availability of manually labeled training data, we propose a synthesized-based learning algorithm for multiple organs that interpolates between two volumes with multiple labels, in which each label corresponds to an organ. Subsequently a retraining of a hierarchical multi organ SSM [14] is performed, using multi-label sets generated by the interpolation.

2 Method

First single-organ level set morphing [8] is described, followed by the proposed multi-organ extension, with the novel overlap prevention approach. Section 2.2 explains the integration of the generated synthetic data into the multi-organ hierarchical SSM.

2.1 Level Set Morphing

Level set morphing is an interpolation algorithm that provides a smooth transition from a source shape to a target shape [8]. Both shape surfaces are described by signed distance functions of the shapes, which is identical to the shape definition from which a level set model is constructed. Level set morphing is a combination of linear and non-linear deformations, as explained in further detail below.

Linear Deformation. Consider an open set, $\Omega_A \subset \mathbf{R}^3$, representing the voxels of a source shape, and a set of voxels for a target shape, $\Omega_B \subset \mathbf{R}^3$. To ensure proper

deformation, the source and target shapes must be aligned. The affine transformation is defined by $\boldsymbol{T}:\mathbf{R}^3 \rightarrow \mathbf{R}^3$, which maps a point $\boldsymbol{x}$ from the source representation Ω_A onto the target representation Ω_B. $\boldsymbol{T}$ is parameterized by translation [t_x, t_y, t_z], rotation [r_x, r_y, r_z] and scaling [s_x, s_y, s_z], hence consisting of nine degrees of freedom. The optimal transformation is acquired by

$$\boldsymbol{T} = \arg\min_{\boldsymbol{T}'} d_{surf}(\boldsymbol{T}' \circ \Omega_A, \Omega_B) \tag{1}$$

where d_{surf} is the surface distance between the two shapes. Note that $\boldsymbol{T}' \circ \Omega_A$ is the source shape, transformed to the target shape.

Non-linear Deformation. To perform non-linear deformation, target Ω_B is mapped onto the source Ω_A. This linear transformation can be expressed by $\Omega'_B = \boldsymbol{T}^{-1} \circ \Omega_B$. Subsequently, using the level set, Ω_A is non-linearly deformed to optimally match Ω'_B. In this deformation process, $\Omega(t)$ is defined as an intermediate shape at time point t, initialized by $\Omega(t)|_{t=0} = \Omega_A$. To steer the deformation progress, the overlap between Ω'_B and $\Omega(t)$ is calculated by

$$M(\Omega'_B, \Omega(t)) = \int_{\Omega(t)} \gamma'_B(\boldsymbol{x}) d\boldsymbol{x} \tag{2}$$

in which γ'_B is an inside-outside function of the target Ω'_B having positive values inside the object and negative outside [8]. Because all the positive values of $\gamma'_B(\boldsymbol{x})$ are elements of Ω'_B and all negative values of $\gamma'_B(\boldsymbol{x})$ are not included in Ω'_B, maximization of eq. (2) results in

$$\Omega'_B = \arg\max_{\Omega} \int_{\Omega} \gamma'_B(\boldsymbol{x}) d\boldsymbol{x} \,. \tag{3}$$

Then, let $\boldsymbol{s}(t)$ be a surface point of $\Omega(t)$, i.e. $\boldsymbol{s}(t) \in \partial\Omega(t)$. During deformation, as shown in [8], the trajectory of every position $\boldsymbol{s}$ can be calculated over time by the Euler-Lagrange equation:

$$\frac{d\boldsymbol{s}}{dt} = \gamma'_B(\boldsymbol{s}(t))\mathbf{N}(\boldsymbol{s}(t)) \quad \forall \boldsymbol{s}(t) \in \partial\Omega(t) \tag{4}$$

with $\mathbf{N}(\boldsymbol{s}(t))$ denoting the outward unit normal of the surface $\partial\Omega(t)$ at $\boldsymbol{s}(t)$, at time point t. Hence, eq. (4) describes the surface motion that minimizes the metric $M(\Omega'_B, \Omega(t))$. The deformation process described above is summarized in Fig. 1.

In this study, the following level set model $\phi(\boldsymbol{x}, t)$ as in [8] is used to obtain the deformation of $\Omega(t)$:

$$\partial\Omega(t) = \{\boldsymbol{x} \mid \phi(\boldsymbol{x}, t) = 0\} \tag{5}$$

which starts at $\partial\Omega_A = \{\boldsymbol{x} \mid \phi(\boldsymbol{x},0) = 0\}$. Given $\nabla\phi(\boldsymbol{x},t) = |\nabla\phi(\boldsymbol{x},t)|\,\mathbf{N}(\boldsymbol{x}(t))$, the surface motion of eq. (4) is modified to

$$\frac{\partial\phi(\boldsymbol{x},t)}{\partial t} = \left|\nabla\phi(\boldsymbol{x},t)\right|\gamma'_B(\boldsymbol{x}) . \tag{6}$$

Then $\phi(\boldsymbol{x}, t)$ is iteratively approximated within a finite grid domain $U \subset \mathbf{R}^3$. Suppose that $u_{i,j,k} \in U$ is a discrete approximation of $\phi(\boldsymbol{x}, t)$, the iterative update function for eq. (6) can be written as:

$$u_{i,j,k} \leftarrow u_{i,j,k} + \Delta t \Delta u_{i,j,k} \tag{7}$$

with i, j and k denoting image coordinates and in which $\Delta u = \partial\phi/\partial t$ is computed by a first order upwind scheme. To enforce a gradual deformation, a large number of small time steps Δt were used in the deformation process, from which a subset was selected such that the average surface-to-surface distance between subsequent deformation stages was equal to d, resulting in a sequence of L intermediate shapes $\Omega(\frac{\ell}{L+1})$ $(\ell = 1, \cdots, L)$.

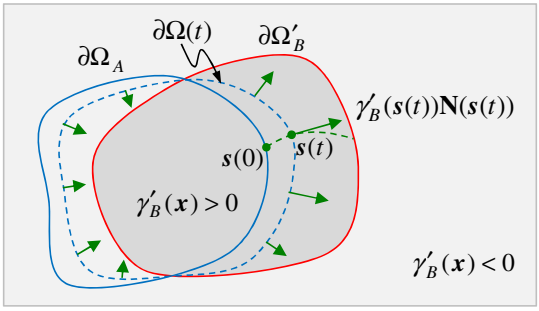


Fig. 1. Illustration of the process of non-linear shape deformation from source shape (blue contour) toward target shape (red contour)

Combination of Linear and Non-linear Deformations. To combine the linear and non-linear deformations, the affine transformation $\boldsymbol{T}$ is divided into L small transformations, jointly corresponding to $\boldsymbol{T}$. Translation, rotation and scaling for such small transformations $\boldsymbol{T}(\alpha)$ are defined $\alpha[t_x, t_y, t_z]$, $\alpha[r_x, r_y, r_z]$ and $\mathbf{1}+\alpha([s_x, s_y, s_z]-\mathbf{1})$ respectively, with $\alpha \in [0,1]$ and $\mathbf{1}$ is a vector with all entries 1. Then, intermediate shapes $\Omega(\frac{1}{L+1}), \cdots, \Omega(\frac{L}{L+1})$ are deformed by $\boldsymbol{T}(\frac{1}{L+1}), \cdots, \boldsymbol{T}(\frac{L}{L+1})$, respectively. Hence, the ℓ-th interpolated shape, following the stepwise deformation trajectory from the source shape to the target shape, is defined by:

$$\Omega^{\ell} = \boldsymbol{T}(\tfrac{\ell}{L+1}) \circ \Omega(\tfrac{\ell}{L+1}) . \tag{8}$$

Extending Level Set Morphing for Multiple Shapes. This paper aims to extend the level set morphing algorithm to be applicable to two sets of multiple organs from different patients. First we apply the level set morphing to each organ independently. To prevent overlaps between the neighboring organs, an energy function is constructed to impose penalties for overlapping volume [2]:

$$E = \sum_{m=1}^{M} \left[\sum_{m'=1,\, m' \neq m}^{M} r_{mm'} \int_{U} (1 - H_{\varepsilon}(\phi_m))(1 - H_{\varepsilon}(\phi_{m'}))dxdydz \right] \tag{9}$$

in which $r_{mm'}$ is a penalty constant for the overlapping organs m and m'. The function ϕ_m is the level set function of organ m and H_{ε} is a regularized Heaviside function. The optimization of ϕ_m is performed over time t and is defined by:

$$\frac{\partial \phi_m}{\partial t} = \delta_{\varepsilon}(\phi_m) \sum_{m'=1,\, m' \neq m}^{M} \left\{ r_{mm'}(1 - H_{\varepsilon}(\phi_{m'})) \right\} \tag{10}$$

in which δ_{ε} is a regularized delta function (the derivative of H_{ε}), and M denotes the number of organs. Using eq. (10), $\{\phi_1, \cdots, \phi_M\}$ are updated simultaneously until they have no overlap with other shapes. This overlap elimination step is applied to all interpolated shapes by eq. (8) and the final shapes are included in the training data of a multi-organ SSM, aiming to boost the performance of the model. This will be further explained in Section 2.2.

2.2 Hierarchical SSM Using Synthesized-Based Learning (S-H-SSMm)

The main aim of this paper is to enhance the performance of multi-organ SSMs, by including a large amount of synthesized training data in the model training data set. To show the benefit of this approach, experiments were performed using a multi-organ hierarchical SSM (H-SSMm) [14]. The H-SSMm trained with additional synthesized training samples, will be referred to as S-H-SSMm through this paper. Because [14] does not provide a detailed description, the remainder of this section will explain the construction of the H-SSMm.

H-SSMm Construction. See Fig 2. Suppose $X_m^n \subset U \quad (m = \{1, \cdots, M\},\ n = \{1, \cdots, N\})$ denotes the integer label of the m-th organ from the n-th patient. After performing generalized Procrustes alignment on all samples, all X_m^n can be decomposed into:

$$X_m^n = Y_m^n \circ \boldsymbol{P}_m^n \tag{11}$$

in which $Y_m^n \subset U$ denotes the shape (including scaling) and $\boldsymbol{P}_m^n : \mathbf{R}^3 \to \mathbf{R}^3$ denotes the pose of the sample, consisting of a translation $\boldsymbol{t}_m^n \in \mathbf{R}^3$ and a rotation $\boldsymbol{r}_m^n \in \mathbf{R}^3$. Subsequently, to analyze the variability across the subjects, (2+M)-fold

PCA is applied: On the concatenated transposed vectors of translation $\boldsymbol{t}_m^n$ $(m=1,\cdots,M)$, on the concatenated transposed vectors of rotation $\boldsymbol{r}_m^n$ $(m=1,\cdots,M)$, and on the signed distance functions $\boldsymbol{\varphi}_m^n \in \mathbf{R}^{|U|}$ of all individual shapes Y_m^n $(n=1,\cdots,N)$, resulting in:

$$\begin{cases} \boldsymbol{t}^n \approx U_T \boldsymbol{\alpha}_T^n + \boldsymbol{\mu}_T \\ \boldsymbol{r}^n \approx U_R \boldsymbol{\alpha}_R^n + \boldsymbol{\mu}_R \\ \boldsymbol{\varphi}_m^n \approx U_{S^m} \boldsymbol{\alpha}_{S^m}^n + \boldsymbol{\mu}_{S^m} \end{cases} \tag{12}$$

with U_T, U_R and U_{S^m} denoting the eigenvectors and $\boldsymbol{\mu}_T$, $\boldsymbol{\mu}_R$ and $\boldsymbol{\mu}_S$ the average representations of the models. After weighting the translation and rotation scores by scalars w_T and w_R respectively, to compensate for differences in the ranges of the distribution, the principal component score vectors are concatenated into vectors:

$$\boldsymbol{b}^n = \left[w_T(\boldsymbol{\alpha}_T^n)^t \quad w_R(\boldsymbol{\alpha}_R^n)^t \quad (\boldsymbol{\alpha}_{S^1}^n)^t \cdots (\boldsymbol{\alpha}_{S^M}^n)^t \right]^t . \tag{13}$$

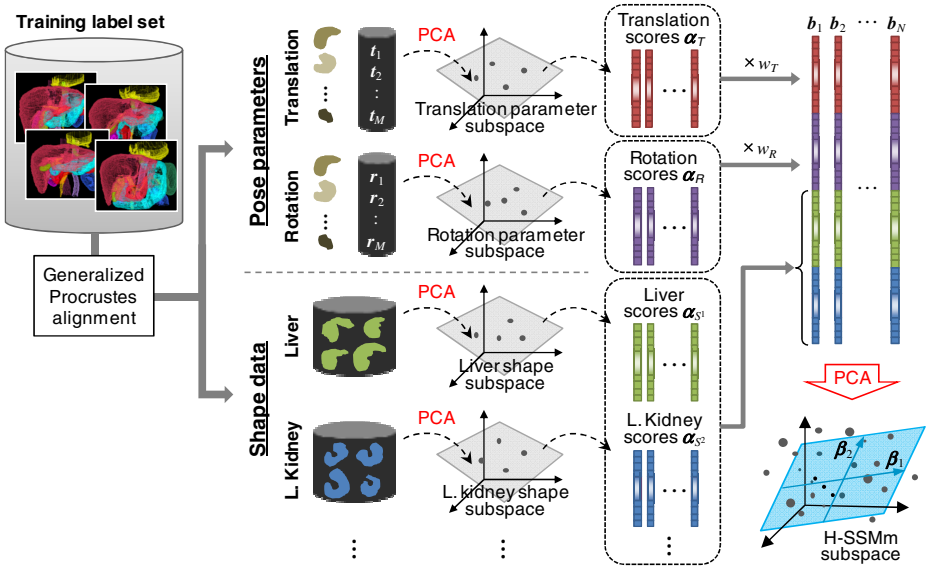


Fig. 2. Hierarchical SSM for multiple organs

Applying PCA on the vectors $\boldsymbol{b}^n$ results in the H-SSMm, with eigenvectors U_H and principal component scores $\boldsymbol{\beta}^n$:

$$\boldsymbol{b}^n \approx U_H \boldsymbol{\beta}^n \tag{14}$$

3 Results

Fourteen abdominal organs were modeled: the liver, spleen, left and right kidneys, heart, gallbladder (GB), pancreas, aorta, inferior vena cava (IVC), portal vein (PV), stomach, esophagus, splenic vein (SV) and superior mesenteric vein (SMV). Fifteen multi-label patient data sets from non-contrast CT volumes served for training the models, eight data sets were used for testing. From the fifteen training sets, 649 multi-organ synthesized data sets were created by the proposed level set morphing with step size d = 2.0 [mm], and added to the training data. Because small organs are more susceptible to changes in volume than larger organs, we used $r_{mm'} = 0.5$ when m corresponded to large organs (liver, spleen, kidneys, heart, pancreas and stomach), and $r_{mm'} = 0.1$ when m corresponded to small organs (GB, aorta, IVC, PV, esophagus., SV and SMV).

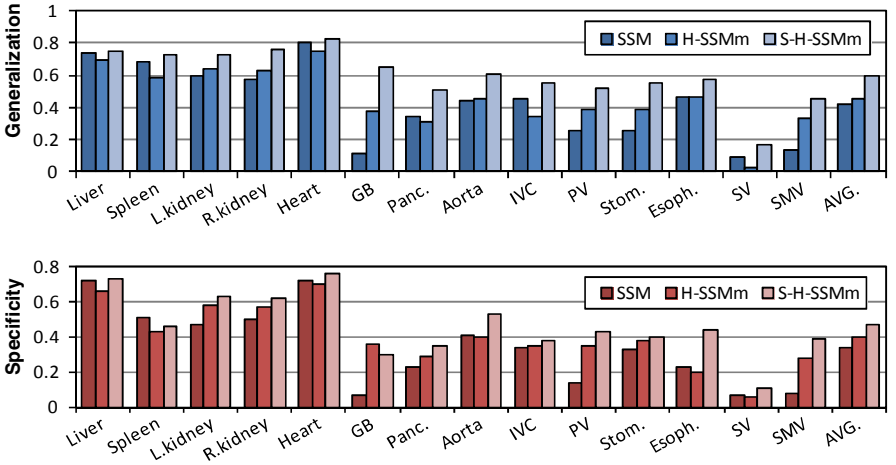


Fig. 3. Average generalization (top) and specificity (bottom) for all organs

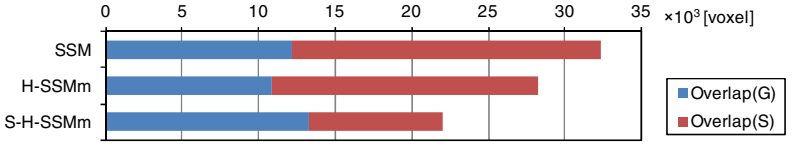


Fig. 4. Average number of voxels containing overlapping organs

Model performances of individual organ SSMs, the H-SSMm and the S-H-SSMm were compared using generalization (the ability to describe unknown shapes) and specificity (the ability to exclude unnatural shapes) [15]. Generalization is defined by the Jaccard Index for the correspondence between a test case and the model approximation of that test case after projection onto, and back-projection from the model

eigenspace. Specificity is a measure for the correspondence between a shape that is created from random model parameters and the closest sample in the test data set. The degree of organ overlap (in number of voxels) is calculated for both the generalization optimization and for the specificity optimization, and are named overlap(G) and overlap(S) respectively. For all three models, the number of principal components was optimized such that the generalization would be maximal.

Fig. 3 shows the results of generalization and specificity for all SSMs for each organ. The degree of overlap is shown in Fig. 4. The generalization, specificity and overlap(S) of the S-H-SSMm were improved by 0.173, 0.124 and 11,459[voxel], with respect to results from individual organ SSMs, and by 0.142, 0.068 and 8,713[voxel] when compared to the H-SSMm. A Wilcoxon signed-rank test (generalization) and a Mann-Whitney *U* test (specificity) showed a statistically significant performance gain when using the S-H-SSMm, both with $p<0.01$. Furthermore, the reduced overlap(G) for the S-H-SSMm proved statistically insignificant ($p>0.1$). The strongly decreased overlap(S) for the S-H-SSMm showed convincing statistical significance ($p<0.01$).

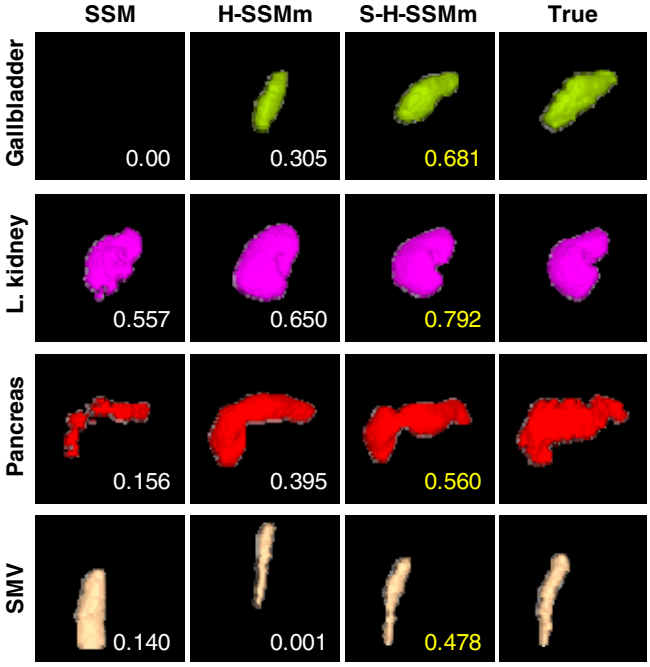


Fig. 5. True shapes and their reconstructed shapes from each model

Fig.5 shows reconstructed shapes for the gallbladder, left kidney and SMV, quantified by the Jaccard Index of the overlap between the reconstructed shape and the true shape. It is apparent that the proposed S-H-SSM produces the best reconstructed shapes, corroborating the results presented in Fig. 3.

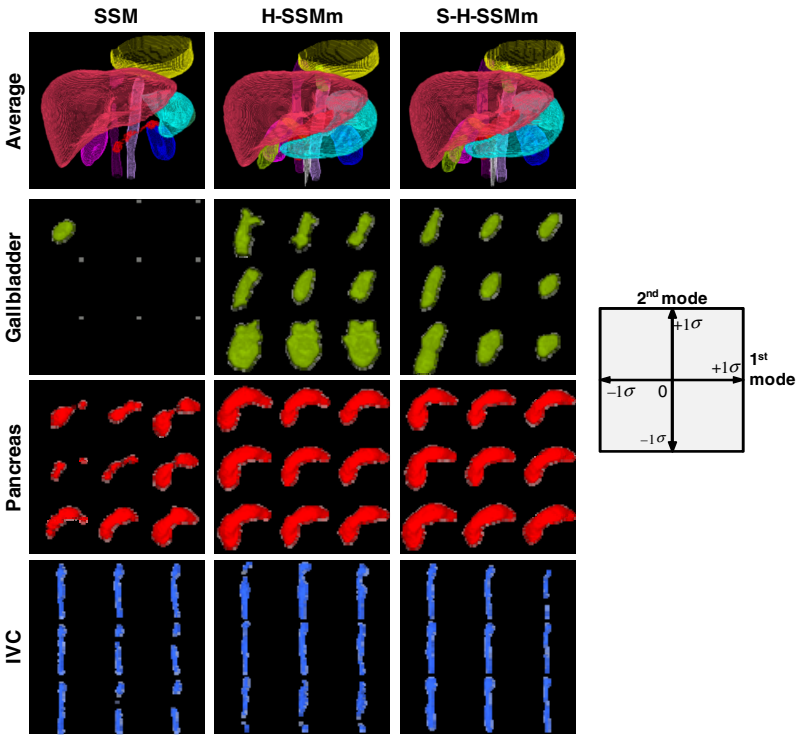


Fig. 6. Average multi-organ representations for all models (top row), and modeled shape representations for gallbladder, pancreas and inferior vena cava (from second to bottom row)

The top row of fig. 6 shows the average multi-organ representation for all compared models. Furhtermore, average shape representations and the variation corresponding to the two largest modes of variation of the (multi-organ) model are shown for the gallbladder, pancreas and IVC. Overall, in terms of topology and shape, the S-H-SSM appears to produce the most natural representation of the organs.

4 Discussion and Conclusion

This paper proposed a multi-organ statistical shape model that was trained on both original and synthesized data. An extension of the level set morphing algorithm was proposed to generate the synthesized training data, by interpolation between two sets of multiple organ labels, while aiming on minimization of organ overlap. Applied to fourteen abdominal organs, using only fifteen multi-organ training data sets, the multi-organ model that was boosted by synthesized data outperformed models that were constructed without synthesized training data. Note that the current work is being done to further validate the findings of this paper. Our future plans include *k*-fold CV on the generalization, specificity and overlaps, as well as on the performance of the multi-organ segmentation based on the proposed multi-organ SSM. We are also

planning to adopt the latest non-linear statistical methods [16] to further improve the performance of the SSMs.

References

1. Park, H., Bland, P.H., Meyer, C.R.: Construction of an abdominal probabilistic atlas and its application in segmentation. IEEE Trans. Med. Imag. 22(4), 483–492 (2003)
2. Shimizu, A., Ohno, R., Ikegami, T., Kobatake, H., Nawano, S.: Segmentation of multiple organs in non-contrast 3D abdominal CT images. IJCARS 2(3-4), 135–142 (2007)
3. Okada, T., Linguraru, M.G., Yoshida, Y., Hori, M., Summers, R.M., Chen, Y.-W., Tomiyama, N., Sato, Y.: Abdominal multi-organ segmentation of CT images based on hierarchical spatial modeling of organ interrelations. In: Yoshida, H., Sakas, G., Linguraru, M.G. (eds.) Abdominal Imaging 2011. LNCS, vol. 7029, pp. 173–180. Springer, Heidelberg (2012)
4. Linguraru, M.G., Pura, J.A., Pamulapati, V., Summers, R.M.: Statistical 4D graphs for multi-organ abdominal segmentation from multiphase CT. Med. Imag. Anal. 16(4), 904–914 (2012)
5. Suzuki, M., Linguraru, M.G., Okada, K.: Multi-organ segmentation with missing organs in abdominal CT images. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 418–425. Springer, Heidelberg (2012)
6. Wolz, R., Chu, C., Misawa, K., Mori, K., Rueckert, D.: Multi-organ abdominal CT segmentation using hierarchically weighted subject-specific atlases. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 10–17. Springer, Heidelberg (2012)
7. Murase, H., Nayar, S.K.: Learning by a generation approach to appearance-based object recognition. In: Proc. of ICPR, vol. 1, pp. 25–29 (1996)
8. Breen, D.E., Whitaker, R.T.: A level set approach for the metamorphosis of solid models. IEEE Trans. on Vis. and Comp. Graph. 7(2), 173–192 (2001)
9. Duta, N., Sonka, M.: Segmentation and interpretation of MR brain images: an improved active shape model. IEEE Trans. Med. Imag. 17(6), 1049–1062 (1988)
10. Tsai, A., Wells, W., Tempany, C., Grimson, E., Willsky, A.: Mutual information in coupled multi-shape model for medical image segmentation. Med. Imag. Anal. 8(4), 429–445 (2004)
11. 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. LNCS, vol. 4191, pp. 955–963. Springer, Heidelberg (2006)
12. Malcolm, J.A., Rathi, Y., Shenton, M.E., Tannenbaum, A.: Label space: A coupled multi-shape representation. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 416–424. Springer, Heidelberg (2008)
13. Changizi, N., Hamarneh, G.: Probabilistic multi-shape representation using an isometric log-ratio mapping. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part III. LNCS, vol. 6363, pp. 563–570. Springer, Heidelberg (2010)
14. Shimizu, A., Nakada, M., Kobatake, H., Nawano, S.: Statistical shape models of multiple organs in an upper abdominal CT volume. International Journal of Computer Assisted Radiology and Surgery 4(supp. 1), S361 (2011)
15. 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)
16. Motai, Y., Yoshida, H.: Principal composite kernel feature analysis: Data-dependent kernel approach. IEEE Trans. Knowl. Data Eng. 25(8), 1863–1875 (2013)

A Survey of Cervix Segmentation Methods in Magnetic Resonance Images

Soumya Ghose[1], Lois Holloway[2], Karen Lim[3], Philip Chan[4], Jacqueline Veera[3], Shalini K. Vinod[3,7], Gary Liney[5], Peter B. Greer[6], and Jason Dowling[1]

[1] CSIRO Computational Informatics, 901/16 UQ Health Science Building, Herston, Qld 4029, Australia
[2] Department of Radiation Oncology, Liverpool Hospital, Elizabeth St Liverpool NSW 2170
Institute of Medical Physics, Sydney University, City Rd, Darlington NSW 2008 and Centre For Medical Radiation Physics, Northfields Ave, Wollongong NSW 2522, University of Wollongong, Australia
[3] Department of Radiation Oncology, Liverpool Hospital, Elizabeth St Liverpool NSW 2170, Australia
[4] Royal Brisbane and Women's Hospital, Butterfield St, Herston QLD 4006, Australia
[5] Ingham Institute for Applied Medical Research, Liverpool Hospital, 1 Campbell Street Liverpool NSW 2170, Australia
[6] Department of Radiation Oncology, Calvary Mater Newcastle Hospital, Edith St, Waratah NSW 2298, Australia and
Department of Physics, University of Newcastle, Callaghan NSW 2308, Australia
[7] University of Western Sydney, Richmond NSW 2753, Australia and
South Western Clinical School, University of NSW, Sydney, Australia

Abstract. Radiotherapy is an effective therapy in the treatment of cervix cancer. However tumor and normal tissue motion and shape deformation of the cervix, the bladder and the rectum over the course of the treatment can limit the efficacy of radiotherapy and safe delivery of the dose. A number of studies have presented the potential benefits of adaptive radiotherapy for cervix cancer with high soft tissue contrast magnetic resonance images. To enable practical implementation of adaptive radiotherapy for the cervix, computer aided segmentation is necessary. Accurate computer aided automatic or semi-automatic segmentation of the cervix is a challenging task due to inter patient shape variation, soft tissue deformation, organ motion, and anatomical changes during the course of the treatment. This article reviews the methods developed for cervix segmentation in magnetic resonance images. The objective of this work is to present different methods for cervix segmentation in the literature highlighting their similarities, differences, strengths and weaknesses.

Keywords: Cervix segmentation methods, registration, statistical shape models, magnetic resonance imaging.

1 Introduction

Cervical cancer is the third most common cancer in women. Worldwide around 530,000 new cases of cervical cancer was diagnosed in 2008 and was responsible

H. Yoshida et al. (Eds.): Abdominal Imaging 2013, LNCS 8198, pp. 290–298, 2013.

for more than 273,000 deaths [1]. Often radiation therapy is used to treat cervical cancer. In external beam radiation therapy, high-energy x-ray beams from multiple directions deposit energy (dose) within the tumor to destroy the cancer cells. The balancing act in designing a radiation treatment plan is to deliver the maximum dose to a patient's tumor while limiting the amount of radiation given to healthy tissues. Therefore identification of the organs and accurate targeting and estimation of dose accumulation in the surrounding tissues of the cervix is absolutely critical.

Image guided radiotherapy aids in delivering high dose to the cervical tumor with minimum damage to the surrounding tissues. However high conformal dose is susceptible to large soft tissue deformation, large inter-fraction organ motion of the cervix and the uterus; increasing the risk of under-dosing the clinically treated volume (CTV). Rectal and bladder filling can significantly displace and deform the cervix and the uterus from their normal positions. Furthermore, Chan et al. [4] recorded large inter-scan movement of the CTV that could be only partially explained by bladder and the rectal filling. Chan et al. suggested adaptive re-planning to compensate for the inter scan movement.

A generous population based CTV to planning target volume (PTV) margin is maintained to account for the geometrical uncertainties resulting in irradiation of healthy tissues. Estimating tumor contour in CT images is difficult and results in significant overestimation of the tumor width, thereby resulting in significant increase in dose delivered to the CTV compared to MRI guided contouring [10,16]. Healthy tissue irradiation could be potentially minimized by using computer aided segmentation methods during treatment planning and online image guidance systems [3].

In recent years Dowling et al. [8] have developed an MRI alone treatment planning and adaptive radiation therapy for the prostate. The soft tissue contrasts are better in MRI enabling better segmentation of the prostate and hence estimation of the CTV. Furthermore such an approach could reduce cost of the treatment and reduce the risk of x-ray radiation for the patient during CT imaging. A pseudo CT image is generated from the MRI of the patient and the segmented contours from MRI are transferred to the pseudo CT images for treatment planning. The dose differences between the pseudo CT and planning CT were quantified and found to be less than 2%.

Adaptive re-planning or MRI guided treatment planning or online image guidance system for radiation therapy would significantly benefit from computer aided segmentation of the cervix. However such methods require a fast, robust and automatic segmentation method to succeed. In recent years, some methods for cervix segmentation in MRI have been reported. This paper presents an up-to-date summary of cervix segmentation methods in MRI. We review the different approaches found in the literature in order to show the similarities and differences and further to extract the advantages and drawbacks from the reviewed algorithms. To have an overall qualitative estimation of the performance of the different methods, we have presented their evaluation metrics and degree of validation.

The outline of the paper is as follows. The state-of-the-art computer-aided cervix segmentation procedures are presented in section 2. In section 3, validation and quantitative evaluation of the segmentation methods are discussed. Similarities, differences, strengths and weaknesses of the adopted approaches are discussed in section 4. Finally, the paper is concluded in section 5.

2 Cervix Segmentation Methods

In this paper we study cervix segmentation methods from the theoretical approach adopted for solving the problem. Computer aided automatic or semi-automatic segmentation of the cervix in MRI is a relatively new domain. From the literature we have selected three papers of Staring et al. [12], Lu et al. [11] and Berendsen et al. [2] from reputed journals like IEEE Transactions in Medical Imaging and Computer Vision and Image Understanding that provide efficient and accurate solutions to a difficult problem. These methods are validated with a large number of datasets and have shown promising results in a clinical setting.

Staring et al. [12] proposed to use the B-spline based non-rigid image registration method for segmenting the CTV, the bladder and the rectum. B-splines consist of a set of control points that can be locally controlled on the image domain. Let $\Omega = \{(x, y, z)|0 \leq x < X, 0 \leq y < Y, 0 \leq z < Z\}$ represent the image domain. The transformation between the moving and fixed images is given by $\mathbf{T}$:$(x, y, z) \mapsto (x', y', z')$, where any point (x, y, z) of the moving image is mapped onto its corresponding point (x', y', z') on the fixed image. Given a mesh of control points on the moving image with a control point defined as $\phi_{i,j,k}$ with uniform spacing of δ mm, the nonrigid transformation $\mathbf{T}$ is defined by B-spline functions as

$$\mathbf{T}(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(w) \phi_{i+l,j+m,k+n} \tag{1}$$

where $i = \lfloor x/\delta \rfloor - 1, j = \lfloor y/\delta \rfloor - 1, k = \lfloor z/\delta \rfloor - 1, u = x/\delta - \lfloor x/\delta \rfloor, v = y/\delta - \lfloor y/\delta \rfloor$. $\lfloor . \rfloor$ and $w = z/\delta - \lfloor z/\delta \rfloor$ is the floor function and B_l represents the l^{th} basis function of the cubic B-spline functions such that $B_0(u) = (1 - u^3)/6, B_1(u) = (3u^3 - 6u^2 + 4)/6, B_2(u) = (-3u^3 + 3u^2 + 3u + 1)/6$ and $B_3(u) = u^3/6$.

The resulting coordinate transformation T_i is applied to the training dataset label L_i to produce the resulting label. Normalized mutual information (NMI) [13] as a similarity metric is maximized to achieve the registration. The NMI is an information theoretic measure that tries to reduce the joint entropy of the images and is given by,

$$\text{NMI} = \zeta_{similarity} = \frac{H(M) + H(F)}{H(M, F)} \tag{2}$$

where $\zeta_{similarity}$ is the similarity measure for B-splines registration that is maximized in the process, $H(M)$ and $H(F)$ are the marginal entropies of the moving

(M) and fixed (F) images respectively, and $H(M,F)$ is the joint entropy of the images. $H(M,F)$ can be written using probability theory as,

$$H(M,F) = -\sum_{m,f} p(m,f)\log\left[p(m,f)\right], \tag{3}$$

where, $p(m,f)$ is the joint probability distribution of the images obtained from their joint histogram.

However unlike the methods of Klein et al. [9] and Chandra et al. [5], Staring et al. [12] proposed to use reliable image feature based registration similar to the work of Toth et al. [14],[15]. Mutual information between the two images was computed on features like the spatial derivative, Hessian and Laplacian. Multi-resolution Gaussian derivatives of these features were computed to make the features invariant to rotation and translation. To reduce the computational complexity involved with the registration procedure principal component analysis (PCA) on the feature set was performed to reduce the dimensionality. Three k-nearest neighbor graphs were constructed from the feature vector and were used to estimate the mutual information between the images. A stochastic gradient descent was adopted to maximize mutual information.

In recent times hybrid methods that combine statistical shape based methods with image registration are being frequently used [15], [8]. One advantage of such an approach is that the shape of the organ under study is explicitly defined in a Gaussian space and by varying the shape parameters new shapes following the Gaussian distribution can be generated. Similar approaches have been adopted for cervix segmentation by Lu et al. [11] and Berendsen et al. [2]. However the shape model significantly differs from one method to another.

Lu et al. [11] used manually segmented planning day MR images to achieve a shape constrained registration to the treatment day image. Shape prior constraints are used to segment the bladder and the uterus. A kernel density estimation approach was adopted for building the shape prior models. Training data generated from manual segmentation are rigidly aligned to minimize pose differences. Each object (the bladder and the uterus) is embedded as the zero level set given by $\Psi_1, \Psi_2, ..., \Psi_n$ to achieve non parametric density estimation of the shape space. To formalize, the kernel density estimation of the shape space is given by,

$$p(\Psi_{S_d}) = \frac{1}{n}\sum_{i=1}^{n} k(D(\Psi_{S_d}, \Psi_i), \sigma) \tag{4}$$

where $D(.)$ is the distance metric in implicit shape space and $k(.,\sigma)$ denotes a Gaussian kernel with kernel size σ, i.e.,

$$k(x,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\frac{-x^2}{2\sigma^2} \tag{5}$$

The distance between the evolving level set and the kernel density estimation of the shape space is minimized to impose shape restriction. The tumor probability map is estimated from a Gaussian mixture model of the intensity of the cervix

region. The Gaussian mixture model enables soft clustering of the normal tissue and the tumor. The non-rigid B-spline based registration is constrained by the tumor probability map and the shape constraints of the bladder and the uterus; to achieve simultaneous segmentation of the tumor, uterus and the bladder.

Berendsen et al. [2] used a B-spline based non-rigid shape-constrained registration for segmentation of the bladder and the CTV. Inter-patient CT to CT affine registration followed by B-spline based non-rigid registration were performed on the bones to align the patients. Intra patient rigid registration between the CT and MRI along with the non-rigid inter patient CT to CT registration was performed to achieve an inter patient MRI registration. The shape model is built from surface meshes of the training images. The point distribution model [6] was built from the aligned 3D points of the surface meshes. PCA of the aligned PDMs identifies the principal modes of shape variations. The statistical shape model is given by,

$$s = \overline{s} + \Phi_s \theta_s \tag{6}$$

where $\overline{s}$ denotes the mean shape, Φ_s contains the first p eigenvectors (obtained from 98% of total variations) of the estimated joint dispersion matrix of shape and θ_s represent the corresponding shape eigenvalues. A penalty function depending on the value of the Mahalanobis distance between the mean mesh with principal modes of variations and the reference mesh is introduced in the non-rigid registration framework to achieve shape prior registration.

3 Validation and Qualitative Evaluations

The performance of cervix segmentation algorithms is evaluated comparing the output of the method with a ground truth (gold standard) obtained from manual delineations of the cervical organs done by experienced radiologists or radiation oncologists. For qualitative evaluation, the obtained segmentation is visually compared with the ground truth while for a quantitative evaluation, an error between the obtained segmentation and the ground truth is numerically presented. The major contour and volume based quantitative error metrics are presented in Table 1. Ideally, a comparison of different state-of-the-art cervix segmentation methods on a public dataset should have been done to evaluate their performances. However, a quantitative comparison is difficult in the absence of public software, datasets and standardized evaluation metrics. Nevertheless, to have an overall qualitative estimate of the functioning of some of the state-of-the-art works in the literature we present the reported results in Tables 2. The index of the table is expanded below.

- The name of the first author has been used as a reference of the paper.
- The segmentation criteria shows the computational technique used in the algorithm.
- The automation (Auto) column specifies the degree of manual interaction that was necessary.
- The measure column refers to the quantitative measures used by the authors to present their obtained results.

Table 1. Evaluation metrics

	Metric	Parameters	Equation	Used by
Contour	Hausdorff distance (HD)	Given a set of finite points $A = \{a_1, a_2, ...a_p\}$ and $B = \{b_1, b_2, ..., b_q\}$	$HD(A,B) = max(h(A,b), h(B,A))$ where $h(A,B) = \max_{a \in A}\left(\min_{b \in B} \|a - b\|\right)$	[2]
	Mean absolute distance (MAD)	Given the distance d_j between each corresponding points $j\ (j = 1, 2, \ldots, N)$ between the algorithmic segmented contours and ground truth.	$MAD = \frac{1}{N}\sum_{j=1}^{N}\left\|d_j\right\|$	[12], [11]
Volume	Dice similarity coefficient (DSC)	TP = True positive, TN = True negative, FP = False positive, and FN = False Negative in voxels	$DSC = \frac{2TP}{(FP+TP)+(TP+FN)}$	[12], [11], [2]
	Sensitivity (SN)		$SN = \frac{TP}{TP+FN}$	[11]
	Specificity (SP)		$SP = \frac{TN}{TN+FP}$	[11]

Table 2. Quantitative evaluations

	Reference	Year	Segmentation Criteria	Auto	Performance		Validation
					Measure	Value	
Cervix Segmentation Methods	Staring [12]	2009	B-spline registration	Manually cropped	CTV DSC	0.86	19 patients×5 weeks=95 datasets
					CTV MAD	3.5 mm	
					Bladder DSC	0.81	
					Bladder MAD	1-2 mm	
					Rectum DSC	0.77	
	Lu [11]	2012	Shape prior levelsets and B-spline registration	Manual segmentation used as prior	Bladder DSC	0.88±0.03	6 patients×6 weeks=36 datasets
					Bladder MAD	1.03±0.16 mm	
					Bladder HD	1.17±0.32 mm	
					Bladder Sensitivity	0.91±2.1	
					Bladder Specificity	0.91±1.28	
					Uterus DSC	0.83±0.04	
					Uterus MAD	1.18±0.24 mm	
					Bladder HD	1.86±0.36 mm	
					Bladder Sensitivity	0.88±2.08	
					Bladder Specificity	0.90±1.28	
					Tumor DSC	0.82±0.03	
	Berendsen [2]	2013	Shape constrained B-spline registration	Manual delineated mask used	CTV mean DSC	0.55	17 patients×5 weeks≈84 datasets
					CTV median DSC	0.57	
					CTV mean HD	36 mm	
					CTV median HD	35 mm	
					Bladder mean DSC	0.67	
					Bladder median DSC	0.73 mm	
					Bladder mean HD	20 mm	
					Bladder median HD	17 mm	

– The validation column gives the number of datasets that were used to validate the algorithm.

4 Discussion

The methods in the literature are primarily focused on bladder, CTV, tumor and rectum segmentation of the cervical images for radiation therapy planning. Primarily these methods depend on B-spline based non-rigid registration for registration of the planning MRI to treatment day MRI. In all methods mutual information is used as a similarity metric for the registration. However,

the registration methods significantly differ from each other from a theoretical perspective. For example Staring et al. [12] uses a multi-feature based mutual information based registration while Lu et al. [11] and Berendsen et al. [2] constrain their registration with shape prior information. It has been observed that feature based registration performs better compared to intensity-based registration as these methods are capable of capturing the underlying edge and texture information from the images [14]. Furthermore the rotational and translational invariance features produce robust registration compared to raw intensities that are frequently corrupted by imaging artifacts like magnetic bias and other noises. The noise is magnified around the edges and hence detecting edges in multi-resolution produces reliable edge information. Hence, the multi-feature mutual information of Staring et al. [12] produces good registration results. However, in later stages of radiotherapy treatment registration of the tumor becomes difficult and a tissue classification framework for tumor segmentation is suggested.

Lu et al. [11] and Berendsen et al. [2] use a shape constraint in their registration framework to improve on segmentation accuracies. However the implementations of the shape constraint significantly differ from one another. While Berendsen et al. uses a classical statistical shape model as proposed by Cootes et al. [6], Lu et al. uses a more recent kernel density estimation of shape representation as proposed by [7]. Shape constrained registration improves the segmentation accuracies as observed in the two papers. However, as Lu et al. points out the shape space is not Gaussian and hence the shape model as proposed by Cootes et al. [6] is not sufficient to handle all deformations in the real world. This is important considering that the organs under study like the bladder, rectum and the uterus may be significantly deformed due to the action of radiation therapy as proposed by [4] and hence the kernel density estimation of the shape space will probably be better for modeling the organs under study.

Tissue classification as proposed by Staring et al. [12] was adopted by Lu et al. [11]. Bayesian estimate of the voxel labels significantly reduces the risk of mis-classification of the healthy tissue and the tumor and when introduced as a probability prior in the registration framework produces good results. However Lu et al. used manual contours as the priors in their registration framework and automatic delineation of the organs under study and the tumor is a challenging task.

5 Conclusions

MRI guided adaptive radiation therapy planning for the cervix could significantly improve cervical cancer treatment. However, success of the procedure is dependent on an automatic cervical organ segmentation algorithm. This paper studies some of the methods developed for cervical image segmentation in MRI. We have highlighted the similarities, differences, strengths and the weaknesses of these methods to enable the reader to make a knowledgeable decision of selecting one method over another. Considering the deformable nature of the organs of interest and radical change of shape over the course of the treatment it would be

difficult to achieve accurate segmentation without incorporating shape information in the model. Furthermore use of appearance information along with edge information detected at multiple scales would provide stability to such models. Shape and appearance space are often considered to be Gaussian and large scale deviation from a mean model for both would produce inaccurate segmentation. Hence it would be useful to consider mixture of Gaussian models for both shape and appearance space to improve accuracy. A tissue classification strategy would be extremely important especially in the advanced stage of the treatment for identifying the gross tumor volume.

References

1. Cervical Cancer Statistics (2013), `http://www.cancerresearchuk.org/cancer-info/cancerstats/types/cervical` (accessed on June 1, 2013)
2. Berendsen, F.F., van der Heide, U.A., Langerak, T.R., Kotte, A.N., Pluim, J.P.: Free-form image registration regularized by a statistical shape model: application to organ segmentation in cervical MR. Computer Vision and Image Understanding (2013)
3. Bondar, L., Hoogeman, M., Mens, J.W., Dhawtal, G., de Pree, I., Ahmad, R., Quint, S., Heijmen, B.: Towards an individualized target motion management for IMRT of cervical cancer based on model-predicted cervix-uterus shape and position. Radiotherapy and Oncology 99, 240–245 (2011)
4. Chan, P., Dinniwell, R., Haider, M.A., Cho, Y.B., Jaffray, D., Lockwood, G., Levin, W., Manchul, L., Fyles, A., Milosevic, M.: Inter- and intrafractional tumor and organ movement in patients with cervical cancer undergoing radiotherapy: A cinematic-MRI point-of-interest study. International Journal of Radiation Oncology Biology Physics 70, 1507–1515 (2008)
5. Chandra, S.S., Dowling, J., Shen, K.K., Raniga, P., Pluim, J.P.W., Greer, P.B., Salvado, O., Fripp, J.: Patient specific prostate segmentation in 3D magnetic resonance images. IEEE Trans. Med. Imaging 31(10), 1955–1964 (2012)
6. Cootes, T.F., Hill, A., Taylor, C.J., Haslam, J.: The Use of Active Shape Model for Locating Structures in Medical Images. Image and Vision Computing 12, 355–366 (1994)
7. Cremers, D., Osher, S., Soatto, S.: Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. International Journal of Computer Vision 69(3), 335–351 (2006)
8. Dowling, J., Lambert, J., Parker, J., Salvado, O., Fripp, J., Wratten, C., Capp, A., Denham, J., Greer, P.: An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy. International Journal of Radiation Oncology Biology Physics 83, e5–e11 (2012)
9. Klein, S., van der Heide, U.A., Lipps, I.M., Vulpen, M.V., Staring, M., Pluim, J.P.W.: Automatic Segmentation of the Prostate in 3D MR Images by Atlas Matching Using Localized Mutual Information. Medical Physics 35, 1407–1417 (2008)
10. Lim, K., Kelly, V., Stewart, J., Xie, J., Cho, Y.B., Moseley, J.B., Brock, K., Fyles, A., Lundin, A., Rehbinder, H., Milosevic, M.: Pelvic radiotherapy for cancer of the cervix: Is what you plan actually what you deliver? International Journal of Radiation Oncology Biology Physics 74, 304–312 (2009)

11. Lu, C., Chelikani, S., Jaffray, D.A., Milosevic, M.F., Staib, L.H., Duncan, J.S.: Simultaneous nonrigid registration, segmentation, and tumor detection in MRI guided cervical cancer radiation therapy. IEEE Trans. Med. Imaging 31(6), 1213–1227 (2012)
12. Staring, M., van der Heide, U.A., Klein, S., Viergever, M.A., Pluim, J.P.W.: Registration of cervical MRI using multifeature mutual information. IEEE Trans. Med. Imaging 28(9), 1412–1421 (2009)
13. Studholme, C., Hill, D.L.J., Hawkes, D.J.: An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition 72(1), 71–86 (1999)
14. Toth, R., Bloch, B.N., Genega, E.M., Rofsky, N.M., Lenkinski, R.E., Rosen, M.A., Kalyanpur, A., Pungavkar, S., Madabhushi, A.: Accurate prostate volume estimation using multifeature active shape models on T2-weighted MR. Academic Radiology 18, 745–754 (2011)
15. Toth, R., Madabhushi, A.: Multifeature landmark-free active appearance models: Application to prostate MRI segmentation. IEEE Trans. Med. Imaging 31(8), 1638–1650 (2012)
16. Viswanathan, A., Dimopoulos, J., Kirisits, C., Berger, D., Potter, R.: Computed tomography versus magnetic resonance imaging-based contouring in cervical cancer brachytherapy: results of a prospective trial and preliminary guidelines for standardized contours. International Journal of Radiation Oncology Biology Physics 68, 491–498 (2007)

Author Index