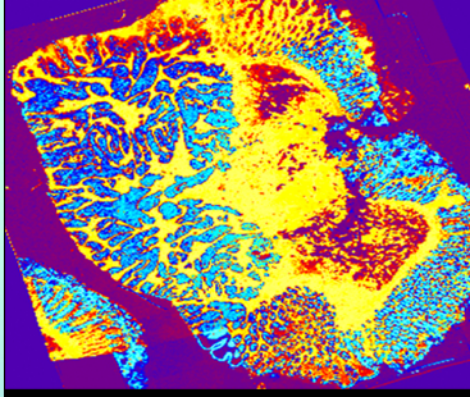
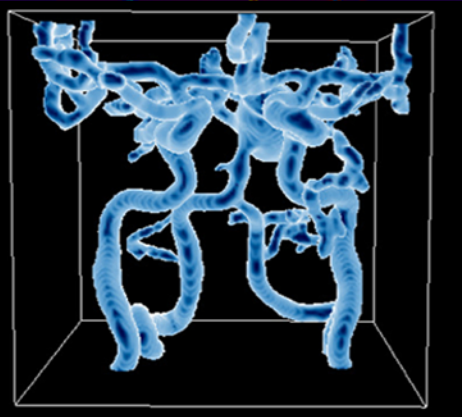
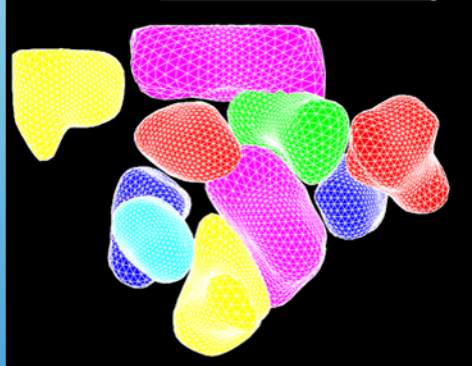
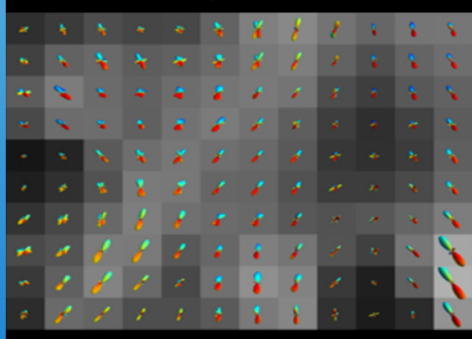
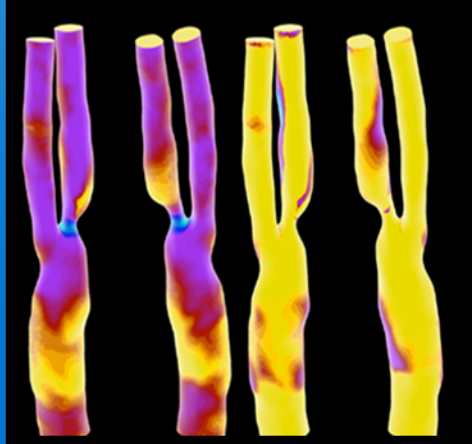
Medical Image Understanding and Analysis 2012

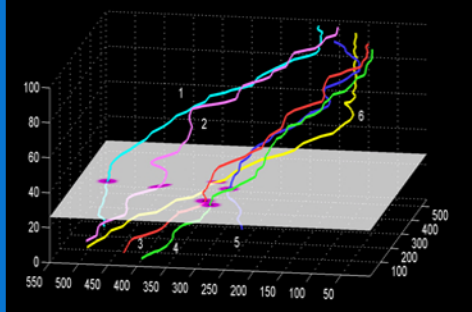












Edited by
Xianghua Xie

Proceedings of the 16th Conference on
Medical Image Understanding and Analysis

9 - 11 July 2012
Swansea, United Kingdom

MEDICAL IMAGE UNDERSTANDING AND ANALYSIS 2012

Xianghua Xie, Editor
Department of Computer Science,
College of Science,
Swansea University
Swansea SA2 8PP
United Kingdom

ISBN: 1-901725-45-6

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Printed and bound in the United Kingdom

www.miua.org.uk

Preface

It is a great pleasure to welcome all participants of the 16th Conference on Medical Image Understanding and Analysis to Swansea. Swansea University is set in rolling parkland overlooking the majestic sweep of Swansea Bay. The campus is a stone's throw from the old fishing village of Mumbles and a short distance to the Maritime Quarter. The University enjoys a prime position overlooking Swansea Bay, the start of the famously dramatic Gower coastline.

MIUA is the principal UK forum for communicating research progress within the community interested in image analysis applied to medicine and related biological science. It is a single-track conference with oral and poster presentations. Authors were asked to submit 6-page technical papers for review by the programme committee. Review papers of up to 8 pages were also welcomed, and we kept the tradition of soliciting short challenge abstract. In total, we received 52 submissions, each of which was reviewed by at least three referees. Based on these reviews, 22 papers were accepted as oral presentation and 16 as posters. We would like to thank all the authors and programme committee members for their efforts.

We are very honoured this year to have four excellent invited talks from Prof. Alison Noble from Oxford University, Prof. Guang-Zhong Yang from Imperial College, London, Prof. Daniel Alexander from University College, London, and Prof. Nassir Navab from Technische Universitat Munchen, Germany. We are grateful to their inspiring contributions.

This is also the first time MIUA has organised a half-day conference tutorial, which is particularly beneficial to research students and early career researchers working in the field. We are grateful to Prof. Nikos Komodakis from the University of Crete, Greece for delivering the conference tutorial.

The Institute of Life Science at Swansea University has kindly offered their venues for hosting the conference. We also would like to thank MIUA steering committee, especially Bill Crum and Neil Thacker, for supporting us organising the event, and a number of people at Swansea University for their efforts in ensuring the conference ran smoothly. Particularly, our thanks are due to Sue Phillips, Ben Daubney, Jonathan Lee Jones, and Jingjing Deng. We thank British Machine Vision Association and WAG Research Institute of Visual Computing for their sponsorship. Finally, we would like to thank all the delegates who attended MIUA 2012 and made it a success.

Xianghua Xie
Swansea University, UK

Sponsors

BMVA

British Machine Vision Association

RIVIC

Research Institute of Visual Computing

With support from

Swansea University
Prifysgol Abertawe

Chair

Xianghua Xie	Swansea University

Steering Committee

Constantino Carlos Reyes Aldasoro	University of Sussex
Ela Claridge	University of Birmingham
Bill Crum	King's College London
Stephen McKenna	University of Dundee
Graeme Penney	King's College London
Nasir Rajpoot	University of Warwick
Neil Thacker	University of Manchester
Xianghua Xie	Swansea University
Reyer Zwiggelaar	Aberystwyth University

Programme Committee

Alison Noble	University of Oxford
Andrew Melbourne	University College London
Andreas Varnavas	King's College London
Andrew King	King's College London
Andrew Bulpitt	University of Leeds
Bill Crum	King's College London
Carlos Reyes-Aldasoro	University of Sussex
David Atkinson	University College London
Dean Barratt	University College London
Daniel Rueckert	Imperial College London
Eddie Edwards	Imperial College London
Ela Claridge	University of Birmingham
Emanuele Trucco	University of Dundee
Fred Labrosse	Aberystwyth University
Graeme Penney	King's College London
Ian Wells	University of Surrey
Jim Graham	University of Manchester
Julia Schnabel	University of Oxford
Kashif Rajpoot	University of Oxford
Majid Mirmehdi	University of Bristol
Nasir Rajpoot	University of Warwick
Neil Thacker	University of Manchester
Paul Aljabar	Imperial College London
Paul Bromiley	University of Manchester
Philip Batchelor	King's College London
Reyer Zwiggelaar	Aberystwyth University
Robert Fisher	University of Edinburgh
Shahrum Nedjati Gilani	King's College London
Stephen McKenna	University of Dundee
Thomas Richardson	King's College London
Tim Cootes	University of Manchester
Tim Nattkemper	Bielefeld University
Xianghua Xie	Swansea University

Local Organisation

Ben Daubney	Swansea University
Jingjing Deng	Swansea University
Jonathan Jones	Swansea University

Table of Contents

Invited Keynote

Conference Tutorial

Classification & Segmentation

Diffusion MRI

Reconstruction, Registration & Multimodality

Image Restoration

Deformable Models & Energy Minimisation

Motion Analysis

Poster Session 1

Poster Session 2

Invited Keynote

Prof. Alison Noble

Oxford University,
United Kingdom

Keynote: Ultrasound Image Analysis: challenges and opportunities

Biography

Professor Alison Noble (FREng) is a Statutory Professor of Biomedical Engineering in the Oxford University Department of Engineering Science and a Fellow of St Hilda's College, Oxford. She is a Senior Member of the IEEE, a Fellow of the IET, and a Fellow of the Royal Academy of Engineering. She is a Director of the Biomedical Image Analysis (BioMedIA) Laboratory, a multi-disciplinary research group working in the area of biomedical imaging and image analysis, a subarea of biomedical engineering. The laboratory is based in the Oxford Institute for Biomedical Engineering (IBME), part of the Department of Engineering Science, which opened in January 2008 and is the hub for biomedical engineering at Oxford.

Professor Noble heads large research activities in cardiovascular image analysis, women's health imaging (obstetrics and perinatal), oncology and smaller activities in image-guided interventions and therapy and cellular image analysis. A particular research strength is ultrasound image analysis (segmentation, registration and RF-signal analysis), with current activities including machine learning for ultrasound segmentation and ultrasound-based biomechanical property estimation. Recent research highlights include the group's publications on fusion echocardiography, multi-modality cardiac image analysis, elasticity and slip imaging for breast cancer diagnosis, and fetal ultrasound image analysis. Much of her group's research is motivated by the need to make more intelligent use of ultrasound information for clinical decision-making and the desire to extract quantitative functional information from spatio-temporal acquisitions (movies of moving objects).

Professor Noble has published around 250 publications. A list of her recent publications can be found at the BioMedIA website Publications webpages. She has supervised 30 PhD students to successful completion, with 17 further DPhil students currently under her sole or co-supervision.

Professor Noble has played a major role in setting up the Biomedical Engineering taught courses and doctoral training programmes at Oxford since 2002. She led setting up and was the first Course Director of the 1yr Oxford MSc in Biomedical Engineering (Director 2006-7,2008-9), also playing a key role in introducing biomedical engineering

into the MEng in Engineering Science during 2002-6. From 2002-8 she was on the management committee of the EPSRC Life Sciences Interface Doctoral Training Centre postgraduate DPhil programme, and the engineering department co-ordinator for that programme. In 2008 she led the successful bid, and is now Director of the £7.1m RCUK Centre for Doctoral Training in Healthcare Innovation, which is pioneering the way to train postgraduate PhD students in healthcare innovation and translational in biomedical engineering. She received a University Teaching Award for her contributions to biomedical engineering teaching and training at the University of Oxford in 2010.

Prof. Guang-Zhong Yang

Imperial College London,
United Kingdom

Keynote: Imaging, Sensing and Robotics for Minimally Invasive Surgery

Reduced patient trauma and recovery time combined with improved precision and dexterity for micro-scale tasks are major benefits of robotically assisted Minimal Invasive Surgery (MIS). This lecture outlines key clinical challenges and research opportunities in developing biomimetic robotic systems integrated with in situ, in vivo imaging and sensing towards the future evolution of medical robotics. The talk will cover the latest developments in fully articulated, bio-inspired robot platforms that facilitate intra-luminal or extra-luminal anatomical curved pathway navigation with integrated sensing and navigation. It addresses key theoretical considerations of bio-mimetic control for intra-operative manipulation under dynamic local/global constraints, as well as the current paradigm shift and clinical demand for bringing cellular and molecular imaging modalities to an in vivo, in situ setting to allow for real-time tissue characterization, functional assessment, and intraoperative guidance. Issues concerning effective, natural human-robot interface are discussed and example implementations for these highly redundant robot platforms based on the concept of perceptual docking are provided.

Biography

Professor Guang-Zhong Yang (FREng, FIEEE, FIET, FAIMBE) is director and co-founder of the Hamlyn Centre for Robotic Surgery, Deputy Chairman of the Institute of Global Health Innovation, Imperial College London, UK. Professor Yang also holds a number of key academic positions at Imperial – he is Director and Founder of the Royal Society/Wolfson Medical Image Computing Laboratory, co-founder of the Wolfson Surgical Technology Laboratory, Chairman of the Centre for Pervasive Sensing. He is a Fellow of the Royal Academy of Engineering, fellow of IEEE, IET, AIMBE and a recipient of the Royal Society Research Merit Award and listed in The Times Eureka 'Top 100' in British Science.

Professor Yang's main research interests are in medical imaging, sensing and robotics. In imaging, he is credited for a number of novel MR phase contrast velocity imaging and computational modelling techniques that have transformed in vivo blood flow quantification and visualization. These include the development of locally focused imaging combined with real-time navigator echoes for resolving respiratory motion for high-

resolution coronary-angiography, as well as MR dynamic flow pressure mapping for which he received the ISMRM I. I Rabi Award. He pioneered the concept of perceptual docking for robotic control, which represents a paradigm shift of learning and knowledge acquisition of motor and perceptual/cognitive behaviour for robotics, as well as the field of Body Sensor Network (BSN) for providing personalized wireless monitoring platforms that are pervasive, intelligent, and context-aware. He has published over 300 peer-reviewed publications, edited over 10 books/conference proceedings, numerous research/best paper awards, and a large research grant portfolio from the UK/EU funding bodies, research charities, and industrial sources.

The Hamlyn Centre (http://www.imperial.ac.uk/robotics) directed by Prof Yang has been established for developing safe, effective and accessible imaging, sensing and robotics technologies that can reshape the future of healthcare for both developing and developed countries. Focusing on technological innovation but with a strong emphasis on clinical translation and direct patient benefit with a global impact, the centre is at the forefront of research in imaging, sensing and robotics for addressing global health challenges associated with demographic, environment, social and economic changes. The Centre plays an active role in international collaboration and outreach activities, as well as in the training of surgeons and engineers in robotic technologies, thereby facilitating a fully integrated clinical approach. It has extensive research, engineering laboratory spaces at the South Kensington campus of Imperial, large pre-clinical facilities at Northwick Park with state-of-the art imaging and surgical equipment for in vivo validation, and comprehensive clinical laboratories at the Surgical Innovation Centre at Imperial College St Mary's Hospital.

In addition to its core research activities, the Centre offers comprehensive PhD and MRes programmes for researchers with a strong technical or clinical background. These programmes are designed to develop cutting edge, disruptive technologies and blue-sky ideas; yet will be appropriate and accessible to both developing and developed countries for addressing different needs of the healthcare challenges with a common ground for technological innovations. Through its endowment fund and close working relationship with industry, government and non-government organisations, the Centre also offers Advanced Fellowships, International Fellowships, Faculty Exchange Programmes, and Research Secondment for specific research projects and technical areas.

Prof. Daniel Alexander

University College London,
United Kingdom

Keynote: Microstructure imaging: next generation diffusion MRI

My talk will focus on the largest current activity of my group: microstructure imaging using magnetic resonance imaging (MRI). Microstructure imaging aims to map histological features of biological tissue (e.g. cell size, shape, density, and arrangement) using non-invasive imaging. Classical histology provides the gold-standard diagnosis for a wide range of diseases, because it is the only way to access such specific information about the cellular architecture of tissue. However, the procedure is invasive and limited to small target areas. Microstructure imaging potentially estimates the same information non-invasively and produces maps over whole organs. It works by fitting a mathematical model of the cellular architecture in each image voxel. Diffusion MRI is a key component of the technique, because the intensity in each image voxel is sensitive to the geometry of cells and membranes within. I will outline the basic ideas and key challenges in realising this technology and describe the current state of the art and on-going work on the topic; see for example (Alexander et al NeuroImage 2010; Drobnjak et al J. Mag. Res. 2010, 2011; Zhang et al NeuroImage 2011, 2012; Panagiotaki et al NeuroImage 2012; Siow et al J. Mag. Res. 2012). Model-based imaging and image analysis is a more general theme of my work, which I shall illustrate with some of the other activities of my group. In particular, I have an emerging strand of work on disease progression modelling, see for example (Fonteijn et al NeuroImage 2012), which my talk will introduce briefly.

Biography

Prof. Daniel Alexander is Professor of Imaging Science in the Department of Computer Science and Centre for Medical Image Computing at UCL. He is an EPSRC Leadership Fellow and co-founder of the EU CONNECT consortium. He leads the Microstructure Imaging Group, which includes 15 researchers many of whom work on joint projects with the Department of Medical Physics, the UCL Institute of Neurology, and the UCL Centre for Advanced Biomedical Imaging. He also leads development and maintenance of the open-source software toolkit, Camino.

Professor Alexander's background is in mathematics and computer science with core expertise in computational modelling and biomedical imaging. His PhD work at UCL was in computer vision, modelling colour image data for autonomous vehicle guidance. His

post-doctoral at the University of Pennsylvania moved him into medical imaging, and he is perhaps best known for his work in brain MRI, most specifically in diffusion MRI. His work on that topic includes seminal contributions to registration of diffusion tensor images, modelling crossing fibres, tractography, pulse sequence and image protocol design, tissue modelling, and microstructure imaging. He also continues to work on more general medical imaging and image analysis problems. Most recently his new strand of work on computational modelling of disease progression focusses on neurological diseases, such as Alzheimer's disease and other dementias, using imaging and non-imaging biomarkers. A theme of his research is a model-based approach to all these problems across a range of scales: from microscopic models of white matter fibre configurations and composition, through whole-brain models of connectivity, up to whole-cohort models of brain atrophy over time. His work has a strong emphasis on translation to basic neuroscience, neuroanatomy, and clinical applications, including multiple sclerosis, dementia, stroke and cancer. The Camino toolkit provides a central mechanism for translation.

Professor Alexander has around 200 peer-reviewed publications. He is associate editor of IEEE Trans. on Medical Imaging (TMI) and PLoS ONE, editorial board member for NeuroImage, and was program chair of BMVC 2009, as well as numerous more specialised workshops and educational sessions on Diffusion MRI. He has graduated 8 doctoral students since starting his academic career at UCL in 2000. He has taught courses on basic programming, data structures, computer vision, image processing, and research methods at advanced MSc level, and starts a new advanced MSc course this year on computational modelling for biomedical imaging. He was program director for the MSc Computer Graphics, Vision and Imaging at UCL until 2008 when his fellowship started.

Prof. Nassir Navab

Technische Universitat Munchen,
Germany

Biography

Prof. Nassir Navab is a full professor and director of the Computer Aided Medical Procedures & Augmented Reality Institute at Technische Universitat Munchen, Germany. He also has a secondary faculty appointment at TU Munchen's Medical School. Prof. Navab has a PhD in computer science from INRIA and the University of Paris XI. In 2006 he was elected as a member of board of director of the MICCAI society. He also served on steering committee of IEEE Symposium on Mixed and Augmented Reality 2000-2008. Prof. Navab acts as associated editor of IEEE Transactions on Medical Imaging and serves on the editorial board of Medical Image Analysis. Prof. Navab is the author of more the 50 US and international patents. He received the Siemens Inventor of the Year Award in 2001 and the SMIT Society Technology Award in 2010. His PhD students have received many prestigious awards including four MICCAI young scientist awards, ISMAR, ISBI, DAGM, VOEC-ICCV and AMDO best paper awards. His research interests include computer vision, augmented reality, computer-aided surgery and medical image registration.

Conference Tutorial

Prof. Nikos Komodakis

University of Crete,
Greece

Biography

Prof. Komodakis is currently serving as an adjunct professor at the Computer Science Department, University of Crete. He is also an affiliated researcher with INRIA-Saclay (the French research institute in informatics and control), working at the computer vision and medical imaging group. Prior to that he had also been a postdoctoral research fellow as well as an adjunct professor at the Applied Mathematics Department of Ecole Centrale de Paris (fellowship awarded by the Agence Nationale de la Recherche), which is one of the top three engineering schools in France (part of the elite of "Grande Ecoles"). In 2007, he was finalist for the prestigious ERCIM Cor Baayen Award for the most promising young researcher in computer science and applied mathematics.

Prof. Komodakis' research interests span various areas of multi-dimensional signal processing and analysis. In particular, he is interested in topics related to biomedical image analysis, computer vision, statistical pattern recognition, image processing, and machine learning. His work has appeared several times in the most reputed conferences and journals from the above fields. Together with his collaborators, he won best paper awards at IPMI 2007 and ISBI 2010. He also serves as a program committee member for a number of top international computer vision and pattern recognition conferences such as ICCV, ECCV and CVPR.

Tutorial: Discrete graphical models for medical image analysis: inference and learning methods

Abstract: Several problems in medical image analysis and computer vision can be formulated using the discrete graphical models framework. Such models are known to provide extreme flexibility and generality, and often lead to state-of-the-art solutions. The two main issues faced by researchers when using graphical models of this type are: (i) Learning: How to estimate the parameters of the model?; and (ii) Inference/optimization: How to find the best assignment for the variables of the model? The main aim of this tutorial will be to thoroughly discuss these two issues, starting from the basics and building up to the state of the art. Furthermore, it will discuss applications of these inference/learning methods to solving fundamental problems in medical imaging and computer vision.

Classification & Segmentation

RanPEC: Random Projections with Ensemble Clustering for Segmentation of Tumor Areas in Breast Histology Images

Adnan Mujahid Khan[1]
amkhan@dcs.warwick.ac.uk
Hesham El-Daly[2]
Hesham.El-Daly@uhcw.nhs.uk
Nasir Rajpoot[1]
nasir@dcs.warwick.ac.uk

[1] Department of Computer Science
University of Warwick
Coventry, UK

[2] University Hospital
Coventry and Warwickshire,
UK

Abstract

Segmentation of areas containing tumor cells in breast histology images is a key task for computer-assisted grading of breast tissue slides. In this paper, we present a fast, unsupervised, and data-independent framework for dimensionality reduction and clustering of high-dimensional data points which we term as Random Projections with Ensemble Clustering (RanPEC). We apply the proposed framework to pixel level classification of tumor vs. non-tumor regions in breast histology images and show that ensemble clustering of random projections of high-dimensional textural feature vectors onto merely 5 dimensions achieves up to 10% higher pixel-level classification accuracy than another state-of-the-art information theoretic method which is both data-dependent and supervised.

1 Introduction

Breast cancer is the leading cancer in women in terms of incidence both in the developed and the developing world. According to the World Health Organization (WHO), the incidence of breast cancer is increasing in the developed world due to increased life expectancy and other factors. Histological grading of breast cancer relies on microscopic examination of Hematoxylin & Eosin (H&E) stained slides and includes: assessment of mitotic count in the most mitotic area, tubule/acinar formation and degree of nuclear pleomorphism over the whole tumor. Recent studies have shown that the proliferation (mitotic) rate provides useful information on prognosis of certain subtypes of breast cancer. This is a highly subjective process by its very nature and consequently leads to inter- and even intra- observer variability. With digital slide scanning technologies becoming ubiquitous in pathology labs around the world, image analysis promises to introduce more objectivity to the practice of histopathology and facilitate its entry into the digital era [5].

Segmentation of areas containing tumor cells in breast histopathology images is a key task for computer-assisted grading of breast tissue slides. Good segmentation of tumor regions can not only highlight areas of the slides consisting of tumor cells, it can also assist in

determining extent of tissue malignancy.Though some algorithms for segmentation of nuclei, quantitative evaluation of nuclear pleomorphism, and grading of lymphocytic infiltration in breast histology images have been proposed in the literature in recent years (see, for example, [2, 5]), tumor segmentation in breast histology images has received relatively less attention.

In this paper, we address the problem of segmentation of tumor regions in a breast histology image using a features based classification approach. The proposed algorithm employs a library of textural features (consisting of just over 200 features), representing each image pixel as a point in a high-dimensional feature space. Due to the so-called *curse of dimensionality*, the high-dimensional feature space becomes computationally intractable and may even contain irrelevant and redundant features which may hinder in achieving high classification accuracy. Recent feature selection and ranking methods, such as the commonly used minimum redundancy maximum relevance (mRMR) [10], employ information theoretic measures to reduce dimensionality of the problem and have demonstrated success in several problem domains. However, major limitations of such approaches include data dependence and the requirement for training the feature selection in a supervised manner. We show that these limitations can be overcome via the proposed Random Projections with Ensemble Clustering (RanPEC) without compromising the segmentation accuracy down to pixel level.

The remainder of this paper is organized as follows. Given a set of features computed for each image pixel, we present a general framework in Section 2 which employs orthogonal random projections with ensemble clustering for assigning a label to each of the image pixels. Section 3 gives some details of the segmentation algorithm, in particular how a library of texture features is computed. Comparative results and discussion are presented in Section 4. The paper concludes with a summary of our results and some directions for future work.

2 Random Projections with Ensemble Clustering

Let $X = \{\mathbf{x}(i,j) \mid (i,j) \in \Omega\}$ denote the set of d-dimensional feature vectors for all pixels in an image $I(i,j), \forall (i,j) \in \Omega$, where Ω denotes the set of all legitimate pixel coordinates for I and $\mathbf{x} \in \mathbb{R}^d$. Suppose now that we reduce the dimensionality of all such vectors to a low-dimensional space $\mathbb{R}^r$ using a linear mapping Φ as follows: $\mathbf{y} = \Phi\mathbf{x}$, where $\mathbf{y} \in \mathbb{R}^r$ and $r << d$ and Φ is a $r \times d$ matrix containing random entries. According to the Johnson-Lindenstrauss Lemma [6], the above mapping can be used to reduce dimensionality of the feature space while approximately preserving the Euclidean distances between pairs of points in the higher d-dimensional space.

One of the major limitations of using random projections for dimensionality reduction and consequently clustering, however, is that the random matrices generated at different runs can produce variable results. Fern *et al.* [4] tackled this issue by generating a similarity matrix from multiple runs of random projections and then using the similarity matrix to drive hierarchical clustering of the data. However, the computational complexity of this approach can make it intractable for use in a large-scale setting. We propose an ensemble clustering approach to address the issue of variability in the results of clustering low dimensional feature data generated by random projections. Let $\{\mathscr{C}_1, \mathscr{C}_2, \ldots, \mathscr{C}_{n_c}\}$ denote the results of clustering the r-dimensional feature data $\mathbf{y}(i,j)$, for all $(i,j) \in \Omega$. In other words, pixel at location (i,j) has n_c labels, where n_c is the number of runs for ensemble clustering. The random projections with ensemble clustering (RanPEC) algorithm for assigning labels to each pixel is given in Algorithm1.

Algorithm 1 Random Projections with Ensemble Clustering (`RanPEC`)

1: **Input**: $X = \{\mathbf{x}(i,j) \mid (i,j) \in \Omega\}$ (where $\mathbf{x} \in \mathbb{R}^d$) the set of high-dimensional feature vectors for all image pixels, r the dimensionality of the lower-dimensional space, n the number of clusters, and n_c the number of runs for ensemble clusters.
2: **Initialization**: Generate random matrices Φ_k, $k = 1,2,\ldots,n_c$, of the order $r \times d$ with matrix entries drawn at random from a normal distribution of zero mean and unit variance.
3: **Orthogonalization**: Use Gram-Schmidt method of orthogonalization to ensure that all rows of Φ_k are orthogonal to each other and have a unit norm. In other words, ensure that $\Phi_k^T \Phi_k$ is an identity matrix, for all $k = 1,2,\ldots,n_c$.
4: **Random Projections**: Project all the feature vectors into r-dimensional space $Y_k = \{\mathbf{y}_k(i,j)\}$ where $\mathbf{y}_k(i,j) = \Phi_k \mathbf{x}(i,j)$ and $\mathbf{y}_k(i,j) \in \mathbb{R}^r$, for all $k = 1,2,\ldots,n_c$ and $(i,j) \in \Omega$.
5: **Ensemble Clustering**: Generate clustering results $\mathscr{C}_k = \{L_k(i,j)\}$ using a clustering method of your choice on the r-dimensional random projections Y_k, for $k = 1,2,\ldots,n_c$ and for all $(i,j) \in \Omega$. Use majority votes in the clustering results to decide the label $L(i,j)$ for image pixel at (i,j) coordinates.
6: **return** $L(i,j)$ for all $(i,j) \in \Omega$.

3 The Segmentation Framework

The RanPEC algorithm described above operates on the set of feature vectors X. In this section, we describe how an input image is pre-processed before computation of feature vectors and application of RanPEC on the feature vectors, followed finally by a set of post-processing operation. An overview of the segmentation framework is shown in Figure 1 with the help of a block diagram. Below we provide a brief description of each of the building blocks, without going into details due to space restrictions.

3.1 Pre-processing

Stain color constancy is one of biggest challenges of H & E staining based on light microscopy. Several factors such as thickness of the tissue section, dye concentration, stain timings, stain reactivity result in variable stain color intensity and contrast. Our pre-processing pipeline consists of stain normalization, background estimation, and edge adaptive smoothing. We used Magee *et al.*'s approach to stain normalization [8]. The background removal was achieved by masking areas containing mostly white pixels. Finally, we converted the stain normalized and background free image into the CIE's La*b* color space and applied anisotropic diffusion to its b* channel in order to remove the inherent camera noise while preserving edges.

3.2 Extraction of Textural Features

We collected a library of frequency domain textural features for each pixel in the image. These consisted of Gabor energy, phase gradients, orientation pyramid, and full wavelet packet features. We used the spatial filter for two-dimensional Gabor function [3] with orientation separation of 30° (i.e., 0°, 30°, 60°, 90°, 120°, and 150°) and 14 scales, resulting in 84 Gabor channel images. Energy of each filter's response at a pixel location was then

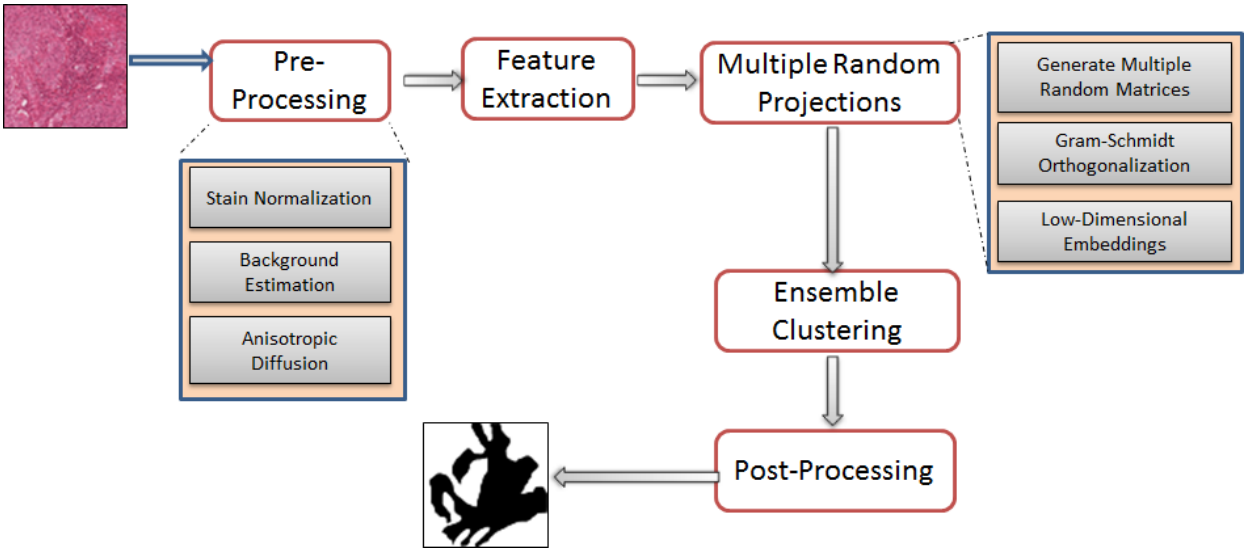


Figure 1: Overview of the proposed tumor segmentation framework.

used as a feature for that filter. Phase information can be used as an important cue in modeling the textural properties of a region. In [9], Murtaza *et al*. used *local frequency* estimates in log-Gabor domain [7] over a range of scales and orientations to yield a signature which uniquely characterizes the texture of a village in satellite images. We computed phase gradient features at 3 scales and 16 orientations to compute 48 filter responses over a window of 15×15 pixels. Next, we used the 3rd level orientation pyramid (OP) features proposed by Wilson & Spann [1, 12], resulting in 21 features. Finally, a set of 64 3-level full wavelet packet features [11] is computed to cater for fine resolution spatial frequency contents in the two texture classes (i.e., tumor and non-tumor). These four sets of features and two proximity features were then concatenated forming a 219-dimensional feature vector per pixel.

3.3 Feature Ranking for Dimensionality Reduction

Feature Ranking (FR) is a family of techniques in which a subset of *relevant* features is used to build a robust learning model that aims to achieve equal, if not better, accuracy of representing high dimensional structures. By removing irrelevant and redundant features from the data, we can improve both the accuracy of learning models and performance in terms of computational resources. Peng *et al.* [10] proposed maximum Relevance Minimum Redundancy (mRMR) feature selection method which employs mutual information to rank features. We compare the performance of mRMR feature selection with the proposed random projections with ensemble clustering (as described in Section 2).

4 Experimental Results and Discussion

Our experimental dataset consisted of digitized images of breast cancer biopsy slides with paraffin embedded sections stained with Hematoxylin and Eosin (H&E) scanned at 40× using an Aperio ScanScope slide scanner. A set of fourteen images was hand segmented by an expert pathologist. We generate all experimental results taking the pathologist's markings as ground truth (GT). All the images are pre-processed in a similar manner, with stain nor-

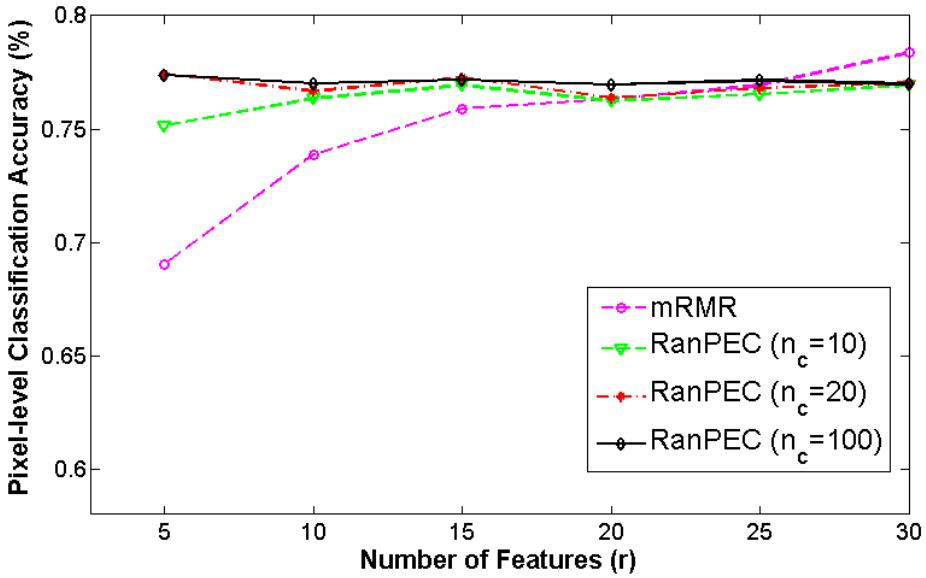


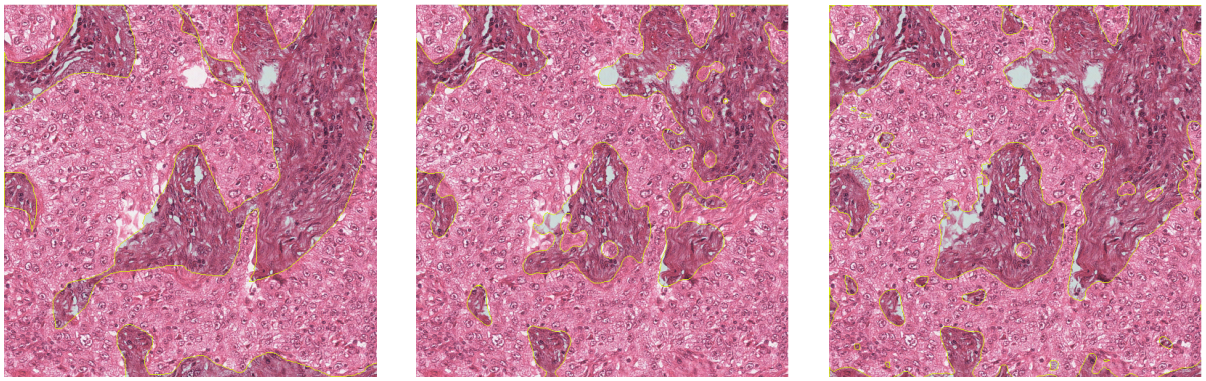
Figure 2: Comparative results of pixel-level classification accuracy (%) versus dimensionality of the feature space for mRMR and RanPEC with n_c=10, 20, and 100.

malization carried out as described in Section 3.1. Background removal is then performed to remove the artifacts caused by staining, fixation and tissue fat. This provides robustness in the subsequent steps of pipeline. As described in Section 3.2, a total of 219 textural and proximity features are calculated for each pixel of the input image I.

Multiple random projections of these textural features are used to generate multiple clustering results from the low-dimensional representation of features using the standard k-means clustering algorithm. A consensus function is then used to combine the partitions generated as a result of multiple random projections into a single partition. A simple majority function is employed on individual cluster labels to produce a single partition. Five replicates of k-means clustering are performed to get a reasonably consistent partitioning.

In order to produce mRMR ranking [10], portion of a subset of GT is chosen as training images. The choice of training images is critical as some of the images have large stromal area and small tumor area while others have vice versa. We ensure that the final training set has approximately similar representation for stromal and tumor areas. Features from all the test images are reordered and k-means clustering is performed. Post-processing is performed on clustering results obtained using both mRMR and RanPEC to eliminate spurious regions and also to merge closely located clusters into larger clusters, producing relatively smooth segmentation results.

Figure 2 presents a quantitative comparison of RanPEC with mRMR. It can be seen from these results that the application of RanPEC with $n_c = 20$ produces quite stable results for almost all values of feature space dimensionality r. Furthermore, the RanPEC results at $r = 5$ generate nearly 10% higher overall pixelwise classification accuracy than mRMR at $r = 5$. Visual results for RanPEC (at $r = 5$)and mRMR (at $r = 65$) as well as the GT are also shown in Figure 3. These results further show that for relatively small dimensionality of the feature space, RanPEC generates superior results than mRMR.



(*a*) (*b*) (*c*)

Figure 3: Visual results of tumor segmentation in a sample image: (*a*) Original image with ground truth (GT) marked non-tumor areas shown in a slightly darker contrast with blue boundaries; (*b*) Results of segmentation with 65-**dimensional** feature space using mRMR (83% accuracy) and (*c*) using RanPEC with 5-**dimensional** feature space and n_c=20 (90% accuracy).

5 Conclusions

In this paper, we addressed the issue of robustness of clustering results in the context of random projections. We proposed a framework for random projections with ensemble clustering and applied it to the segmentation of tumor areas in breast cancer histology images. We showed that the proposed framework RanPEC preserves the Euclidean distance between points in high-dimensional spaces in a robust manner. For our application of tumor segmentation, reasonably high accuracy was achieved using only 5 dimensional feature space.

6 Acknowledgments

The authors would like to thank Warwick Postgraduate Research Scholarship (WPRS) and Department of Computer Science, University of Warwick for funding the research work of AMK. The images used in this paper are part of MITOS dataset, a dataset setup for ANR French project MICO.

References

[1] C.C.R. Aldasoro and A. Bhalerao. Volumetric texture segmentation by discriminant feature selection and multiresolution classification. *Medical Imaging, IEEE Transactions on*, 26(1):1–14, 2007.

[2] A.N. Basavanhally, S. Ganesan, S. Agner, J.P. Monaco, M.D. Feldman, J.E. Tomaszewski, G. Bhanot, and A. Madabhushi. Computerized image-based detection and grading of lymphocytic infiltration in her2+ breast cancer histopathology. *Biomedical Engineering, IEEE Transactions on*, 57(3):642–653, 2010.

[3] J.G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and

orientation optimized by two-dimensional visual cortical filters. *Optical Society of America, Journal, A: Optics and Image Science*, 2:1160–1169, 1985.

[4] X.Z. Fern and C.E. Brodley. Random projection for high dimensional data clustering: A cluster ensemble approach. In *International Conference on Machine Learning*, volume 20, page 186, 2003.

[5] M.N. Gurcan, L.E. Boucheron, A. Can, A. Madabhushi, N.M. Rajpoot, and B. Yener. Histopathological image analysis: A review. *Biomedical Engineering, IEEE Reviews in*, 2:147–171, 2009.

[6] W.B. Johnson and J. Lindenstrauss. Extensions of lipschitz mappings into a hilbert space. *Contemporary mathematics*, 26(189-206):1–1, 1984.

[7] H. Knutsson. *Filtering and reconstruction in image processing*. Linköping University Electronic Press,, 1982.

[8] D. Magee, D. Treanor, P. Chomphuwiset, and P. Quirke. Context aware colour classification in digital microscopy. In *Proc. Medical Image Understanding and Analysis*, pages 1–5. Citeseer, 2010.

[9] K. Murtaza, S. Khan, and N. Rajpoot. Villagefinder: Segmentation of nucleated villages in satellite imagery. *British Mission Vision Conference*, 2009.

[10] H. Peng, F. Long, and C. Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1226–1238, 2005.

[11] N.M. Rajpoot. Texture classification using discriminant wavelet packet subbands. In *Circuits and Systems, 2002. MWSCAS-2002. The 2002 45th Midwest Symposium on*, volume 3, pages III–300. IEEE, 2002.

[12] R. Wilson and M. Spann. *Image segmentation and uncertainty*. John Wiley & Sons, Inc., 1988.

Tumour segmentation in breast tissue microarray images using spin-context

Shazia Akbar[1]
shaziaakbar@computing.dundee.ac.uk
Telmo Amaral[2]
telmoamaral@sapo.pt
Stephen J. McKenna[1]
stephen@computing.dundee.ac.uk
Alastair Thompson[3]
a.m.thompson@dundee.ac.uk
Lee Jordan[4]
lee.jordan@nhs.net

[1] School of Computing
University of Dundee
[2] Institute of Systems and Robotics
University of Coimbra
[3] Dundee Cancer Centre
Ninewells Hospital, Dundee
[4] Pathology
Ninewells Hospital, Dundee

Abstract

A method for automatic segmentation of tumour regions in breast histopathology images is described. It uses auto-context to label pixels based on local image features and contextual label probabilities. We propose spin-context to compute context features that are invariant under image rotation. Quantitative evaluation is reported using spots stained for estrogen receptor. The use of context resulted in improved segmentation.

1 Introduction

Tissue microarrays (TMAs) are used extensively to analyse various types of cancer for molecular and protein markers. Annotation software for breast tissue histopathology images often requires a pathologist to partially annotate some tissue components in order for the software to then analyse a whole mount slide. When applied to TMA spots, typically 0.6*mm* in diameter, regions are often mislabelled due to lack of context. In this paper a method is reported that probabilistically classifies pixels as non-tumour or tumour (invasive or in-situ carcinoma). We propose a distribution-based auto-context descriptor called spin-context and report results on estrogen-receptor stained TMA images ('spots') by comparison to manual segmentation performed by a pathologist.

2 Related Work

Related work on tumour segmentation includes that of Chomphuwiset *et al.* [1], who used Hough transform-based techniques to detect cell nuclei in liver histopathology images. Chomphuwiset *et al.* also integrate a Bayesian network to combine random forest classification results, obtained from texture features, with context information from nearby nuclei and

regions. Wang *et al.* [9] proposed a method for segmentation of tumour, stroma and inflammatory cells in TMA images using tissue architecture extraction and a tumour texture learning model. Tissue architecture extraction consisted of a stain separation method and an unsupervised multistage entropy-based segmentation method. Tumour texture learning consisted of a Markov random field image segmentation system.

Auto-context has been used for medical image segmentation. Morra *et al.* [5] used AdaBoost with auto-context to segment hippocampus in 3D structural MRI. Tu *et al.* [8] used auto-context to segment multiple structures in brain MRI. Tao *et al.* [7] used Gaussian mixtures with simplified auto-context to segment ground glass nodules in 3D lung CT data. Montillo *et al.* [4] segmented structures such as aorta, pelvis, and lungs in 3D CT data, proposing an extension of decision forest classifiers that incorporates semantic context in a manner similar to auto-context. Jurrus *et al.* [2] described an auto-context method to detect membranes in electron micrographs. All of the above were not distribution-based descriptors and, appropriately for those applications, were not invariant under image rotation. To the best of our knowledge, auto-context has not been applied to segmentation of 2D histopathology images.

3 Method

We address locating carcinoma (invasive or in situ) in TMA spots. This is formulated as classifying each location on a grid as being tumour or non-tumour. The image patch around each location is characterised using local features extracted at full resolution, specifically differential invariants up to 2nd order [6] and intensity spin image features [3]. We propose a method called *spin-context* which incorporates context in a rotationally invariant fashion, as the rotation of the tissue in histopathology images is arbitrary. Before describing spin-context we briefly describe auto-context classification [8] and spin image features [3].

Intensity-domain spin image features were proposed for texture representation [3]. A spin feature encodes the distribution of brightness values within a circular support region centred at a location, $\mathbf{x}_0$, using a histogram representation that is invariant under image rotation. The contribution of a pixel $\mathbf{x}$ depends on its intensity value, $I(\mathbf{x})$, and its distance from $\mathbf{x}_0$, $||\mathbf{x}-\mathbf{x}_0||$, as shown in Equation (1). α and β are parameters that determine bin size in the two-dimensional 'soft' histogram, H, where each bin is indexed by the radial distance interval, d, and intensity interval, i.

$$H_{(d,i)} = \sum_x \exp\left(-\frac{(||\mathbf{x}-\mathbf{x}_0|| - d)^2}{2\alpha^2} - \frac{|I(\mathbf{x}) - i|^2}{2\beta^2}\right) \tag{1}$$

The auto-context method, proposed by Tu and Bai in 2009 [8], is an iterative pixel labelling technique, in which some of the label probabilities output at a given iteration are used as contextual data that are concatenated with local image features to form the input vector for the following iteration. This technique is formally described in Algorithm 1, where, for pixel n, t^n is the ground truth value (0 for background, 1 for tumour), $\mathbf{z}^n$ is the local feature vector, $\mathbf{c}^n$ is the context feature vector, and y^n is the predicted probability of being tumour. Classifier models are denoted by ϕ, and $\mathscr{N}$ is the index set for all pixels (training and testing). In a cross-validation experiment (e.g. ten-fold), at the start of each fold p we (1) initialise all pixel probabilities with a uniform prior; and (2) identify the set of training pixel indices. Then, at each auto-context iteration j, we (a) compute a context feature vector from the probability values predicted at the previous iteration, for all pixels; (b) train a classifier

Algorithm 1 Auto-context.

1. initialise $y^n_{p,0}$, for $n \in \mathcal{N}$
2. identify $\mathcal{N}_p^{train} \subset \mathcal{N}$
3. for each iteration j
   (a) compute $\mathbf{c}^n_{p,j}$ from $y^n_{p,j-1}$, for $n \in \mathcal{N}$
   (b) train $\phi_{p,j}$ to predict t^n from $\mathbf{z}^n$ and $\mathbf{c}^n_{p,j}$, for $n \in \mathcal{N}_p^{train}$
   (c) use $\phi_{p,j}$ to predict $y^n_{p,j}$ from $\mathbf{z}^n$ and $\mathbf{c}^n_{p,j}$, for $n \in \mathcal{N}$

to predict truth values from local and context feature vectors, for training pixels; and (c) use the trained classifier to predict new probability values from local and context feature vectors, for all pixels.

3.1 Spin-context

Tu and Bai used a star shaped 'stencil' to select context location points around the pixel being classified. The resulting context features from this stencil were not invariant under image rotation. By using an alternative *spin-context*, we can compute rotationally invariant context features for a given grid location from label probability values within a circular support region. Spin-context is extracted analogously to intensity spin features, computing a two-dimensional 'soft' histogram reflecting the distribution of probabilities within the support region, with rows representing probability intervals and columns representing radial distance intervals. Figure 1 illustrates spin-context for a given support region.

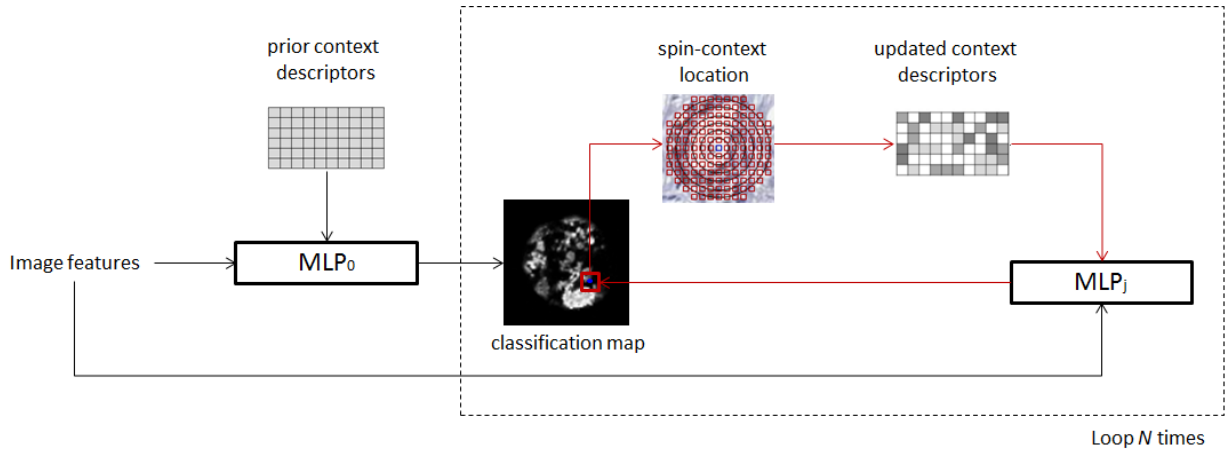


Figure 1: Spin-context constructs context descriptions for a point to be classified (the blue dot) by applying a circular support region centred on that location. The resulting classification map produced by the MLP classifier updates context descriptors iteratively.

The spin-context descriptor allows clutter outside the tissue spot's boundary to be disregarded while considering only data within the spot region. Figure 2 illustrates the advantage of using spin-context to produce a more accurate representation of context information around the boundaries of the spot. The use of boundary information prior to context extraction allows the contributions of out-of-boundary points towards the two-dimensional spin

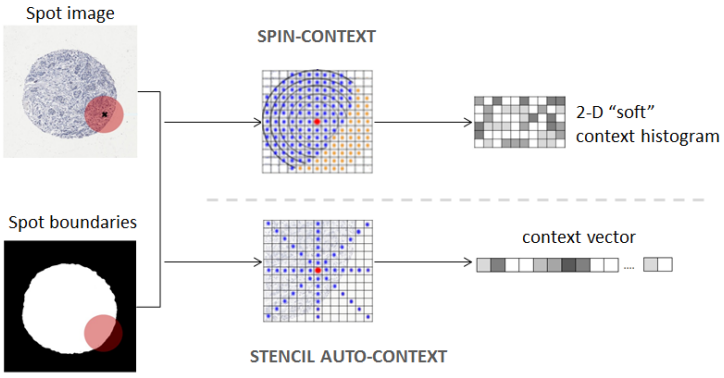


Figure 2: A binary mask is used to ignore the contributions of pixels outside the spot's boundary towards the spin histogram. Stencil-context, however, corresponds to label probability values at all locations lying on a star-shaped stencil, regardless of spot boundaries.

histogram to be ignored. In doing so, not only is context information accurate for the current iteration but subsequent iterations also reflect on accurate information extracted from the spot region. The stencil-context descriptor, not being distribution-based, does not allow this level of flexibility to be maintained, resulting in background interference.

4 Experiments

TMA spots were subjected to nuclear staining for estrogen receptor (ER). Spot images were 4000×4000 pixels. Data consisted of 64 images, 32 of which contained tumour regions annotated by a highly experienced pathologist and 32 were confirmed to contain only healthy tissue. Example pathologist annotations are shown in Figure 4.

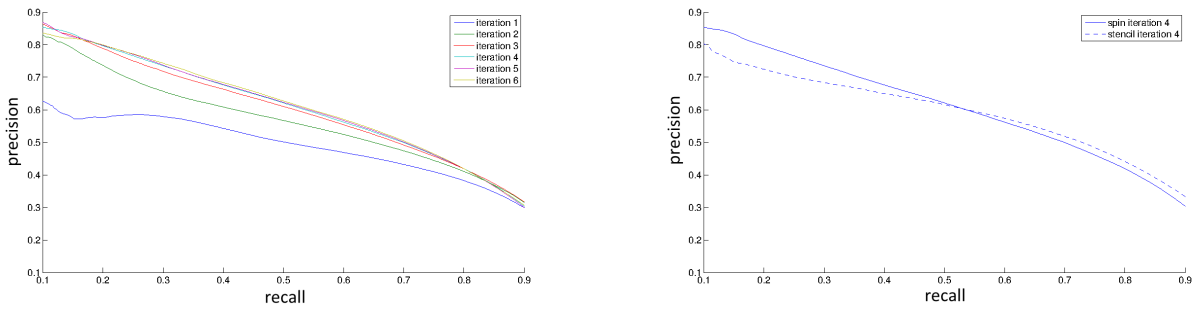


Figure 3: Precision-recall curves for tumour localisation. (a) Effect of six spin-context iterations on MLP classification. (b) Comparison of stencil-context and spin-context for iteration 4 on MLP classification.

Tumour labelling was evaluated using ten-fold cross-validation on the 64 spots. Multilayer percepton (MLP) classifiers had five hidden units, a regularisation constant of 0.1 and used scaled conjugate gradient optimisation. Local and context features were computed at points on a 76×76 grid (a grid step of 50 pixels). Differential invariant features were computed at three scales using a Gaussian pyramid and filters with a standard deviation of 8 pixels. Intensity spin local features were computed at two scales (again using a Gaussian

pyramid) with a circular support region with a radius of 50 pixels. Spin-context used a circular support region with a radius of six grid points. We also tried auto-context (non-rotationally invariant context) using a stencil in which neighbouring grid points within a radius of six grid spacings in each of the eight cardinal and inter-cardinal compass directions were used as context. Labellings obtained were compared to ground-truth segmentations provided by the pathologist to compute precision-recall curves.

5 Results

The precision-recall curve in Figure 3(a) displays the results obtained for six spin-context iterations. All six iterations of spin-context improved results, however this improvement was noticeable in the initial four iterations, whereas the subsequent two iterations improved results marginally. Figure 3(b) compares spin-context with stencil based auto-context. At lower recall values spin-context was superior. In both cases MLP classifiers were used and six context iterations were executed.

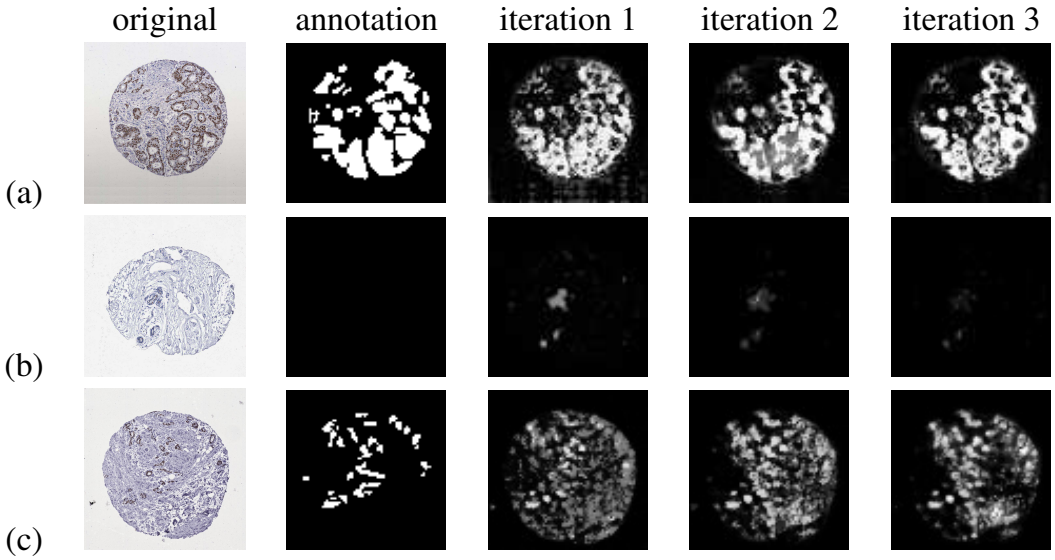


Figure 4: Tumour location probabilities obtained by spin-context. Shown for each TMA spot are the pathologist's annotation, the labelling obtained using local image features (iteration 1) and labellings obtained after incorporating label context (iteration 2 and iteration 3). (a) shows invasive cancer labelled largely in agreement with the pathologist. (b) shows healthy tissue. (c) shows one of the worst results obtained.

Figure 4 shows three spots, two containing tumour and one not containing tumour, along with their expert annotations and the outputs of the spin-context method. In Figure 4(a), posterior probabilities within tumour regions increased at each iteration, so that after the final iteration they were above 0.9 for most tumour pixels. In Figure 4(b), non-zero probabilities occur within regions of normal tissue at the first iteration; however, their values decreased after further iterations, so that a binarisation of the labelling would result in an almost entirely empty (i.e. correct) output. Figure 4(c) shows a case of unsuccessful labelling. Initially, the entire spot region was lightly detected and gradually removed from the left after subsequent iterations. This may have been due to the higher density of tissue which is commonly found in tumour regions in training data. This suggests the need for more training examples of this type.

6 Conclusion

A method for tumour segmentation was presented incorporating rotationally invariant context features. It was validated against manual annotations provided by a pathologist. Figure 4 shows how spin-context can be useful to pathology research in locating tumour regions.

Acknowledgements

This work was supported by the Chief Scientist Office (grant no. CZB/4/761); and the Engineering and Physical Sciences Research Council (DTA grant).

References

[1] P. Chomphuwiset, D. Magee, R. Boyle, and D. Treanor. Context-based classification of cell nuclei and tissue regions in liver histopathology. In *Proc. Medical Image Understanding and Analysis*, 2011.

[2] E. Jurrus, A.R.C. Paiva, S. Watanabe, J.R. Anderson, B.W. Jones, R.T. Whitaker, E.M. Jorgensen, R.E. Marc, and T. Tasdizen. Detection of neuron membranes in electron microscopy images using a serial neural network architecture. *Medical Image Analysis*, 14(6):770–783, 2010.

[3] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture representation using local affine regions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1265–1278, 2005.

[4] A. Montillo, J. Shotton, J. Winn, J. Iglesias, D. Metaxas, and A. Criminisi. Entangled decision forests and their application for semantic segmentation of CT images. In *Information Processing in Medical Imaging*, pages 184–196. Springer, 2011.

[5] J.H. Morra, Z. Tu, L.G. Apostolova, A.E. Green, C. Avedissian, S.K. Madsen, N. Parikshak, X. Hua, A.W. Toga, C.R. Jack Jr, et al. Validation of a fully automated 3D hippocampal segmentation method using subjects with alzheimer's disease mild cognitive impairment, and elderly controls. *Neuroimage*, 43(1):59–68, 2008.

[6] C. Schmid and R. Mohr. Matching by local invariants. Technical Report RR-2644, INRIA, 1995.

[7] Y. Tao, L. Lu, M. Dewan, A. Chen, J. Corso, J. Xuan, M. Salganicoff, and A. Krishnan. Multi-level ground glass nodule detection and segmentation in CT lung images. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pages 715–723. Springer, 2009.

[8] Z. Tu and X. Bai. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(10):1744–1757, 2010.

[9] C. Wang, D. Fennel, I. Paul, K. Savage, and P. Hamilton. Robust automated tumour segmentation on histological and immunohistochemical tissue images. *PLoS ONE*, 6 (2):e15818, 2011.

Nakagami-Based Choroid Plexus Detection in Fetal Ultrasound Images using AdaBoost

Ana I. L. Namburete
ana.namburete@eng.ox.ac.uk
Bahbibi Rahmatullah
bahbibi.rahmatullah@eng.ox.ac.uk
J. Alison Noble
alison.noble@eng.ox.ac.uk

BioMedIA Laboratory
Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford
Oxford, UK

Abstract

In obstetrical ultrasound examinations, the appropriate fetal biometry plane is identified by the presence and absence of key anatomical landmarks within the image [3]. In the 20 to 22 weeks of gestation, the transthalamic plane becomes of particular importance, as it is the stage at which early signs of trisomy 18 can be detected by the presence of cysts in the choroid plexus (CP) [5]. We propose a method to detect the presence of the CP in 2D ultrasound images of the fetal brain using an AdaBoost learning algorithm. We compared the performance of the detection using three different feature sets: intensity-based, and empirically-fit Nakagami μ and ω parameter features. We found that the accuracy of the Nakagami ω parameter had the highest detection accuracy (72.73%).

1 Introduction

Obstetrical ultrasound examinations form an integral part of standard prenatal monitoring. It serves for the purpose of pregnancy dating, fetal growth monitoring, and detection of abnormalities throughout gestation. During the examination, the ultrasonographer collects scans from the fetal head, abdomen, and thigh such that the standard fetal biometric measurements can be obtained [3]. These measurements are then plotted on population-based growth charts to assess the normality of fetal growth. Our interest is assessing fetal development from analysis of the fetal brain. In the standard neurosonography examination, an axial plane (the *transthalamic* plane) is used for assessing the fetal head, as it is believed to contain the significant anatomical landmarks (Fig. 1), yet it is the most challenging plane to locate [3]. The biometric measurements of the fetal head (namely the biparietal diameter (BPD) and occipital-frontal diameter (OFD)) are acquired in this plane and are plotted on the growth curves to determine gestational age of the fetus [3].

1.1 Transthalamic Plane and the Choroid Plexus

In the transthalamic plane, the choroid plexus appears as an echogenic glomular region in ultrasound images, and it has slight variations in size and shape across gestational age [1]. The basic scan is performed at 20 to 22 weeks [3] in which the appearance of the CP is of

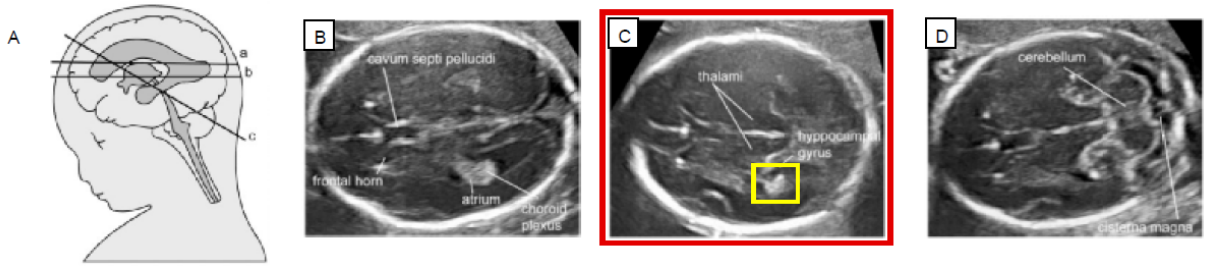


Figure 1: Three axial ultrasound planes used in standard fetal neurosonography examination (A): transventricular plane (B), transthalamic plane with yellow box around the choroid plexus(C), transcerebellar place (D). (Adapted from [3])

particular importance in determining the presence of cysts within the CP (which are indicative of early signs of trisomy 18) [5]. The current practice in an obstetrical examination is for the ultrasonographer to determine whether a suitable transthalamic plane has been selected for biometrical measurements by looking for a number of visual cues including texture, size, shape, and boundary characteristics, which means that the diagnostic decisions are influenced by the experience of the ultrasonographer.

Another difficulty with the detection of structures in ultrasound images is their susceptibility to data drop-out noise commonly described as speckle. Particularly if the plane is selected based on the appearance characteristics of the object of interest within the image, it seems appropriate to design an automated classifier that depends on the backscatter properties of the choroid plexus. Recent studies have shown that statistical processing is beneficial for characterizing tissues in ultrasound images, namely the Nakagami distribution and its parameters [7]. In the current study, we have empirically fit a Nakagami distribution to ultrasound image regions and used the information in an object detection framework.

1.2 Related Work

In the work of Rahmatullah et al. [4], a machine learning technique was proposed for checking that a given image plane has key structures in it for biometric measurements. They used an AdaBoost object detection framework which was robust to size, shape, and position variations of the object of interest in an obstetrical ultrasound image, which suggests it may be suitable for the detection of the choroid plexus.

Traditionally, the AdaBoost-based detection framework has been implemented using intensity-based features. In this study, we instead use an empirical model of ultrasound image formation to account for the speckle within the image and to determine whether the scatter profile of a tissue can be used as a form of tissue characterization. We used the local-based estimation of Nakagami parameters proposed by Wachinger et al [7]. We compared the performance of features extracted from the statistical parameter plots of the images to those of intensity images in the detection of an object in an ultrasound image.

To the best of our knowledge, the automated detection of the choroid plexus has not been attempted before. Thus, the contributions of this study are *a*) to present a method for detecting an anatomical landmark in the standard transthalamic fetal brain scan; *b*) to introduce a novel feature set (Nakagami) in the machine learning framework for object detection in ultrasound images; *c*) to determine the performance of the Nakagami-based feature set in comparison with the currently-used intensity-based ultrasound object detection.

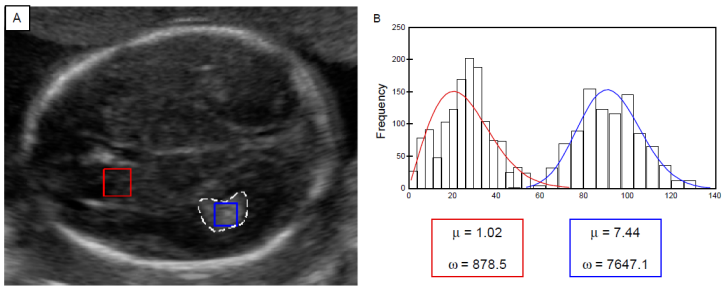


Figure 2: Empirically-fitted histograms and Nakagami parameters for rectangular regions inside the choroid plexus (in blue) and in a background region (in red). The choroid plexus is outlined by a white dotted line.

2 Methods

Two-dimensional ultrasound images of the standard transthalamic plane of the fetal brain were randomly selected from the ongoing Intergrowth-21st study database. For this work, the images were acquired between 19^{+6} to 22^{+6} weeks of gestation. All women were screened based on a set of criteria that define them as having low risk of impairment of fetal growth or fetal abnormality. The images were all acquired using a 2D linear probe (Phillips HD9) at a 2-5MHz wave frequency and saved in a DICOM format of size 1024×728 pixels.

2.1 Nakagami Distribution Approximation

Tissue characterization and detection in ultrasound images is particularly challenging due to the presence of speckle and inhomogeneities within the images. In recent years, segmentation and classification studies have investigated a multitude of distributions for modelling ultrasound scattering for varying amounts of scatterer per resolution cell, with the Rayleigh distribution being the most commonly applied. The Nakagami distribution (which has the properties of an incomplete Gamma function) is of particular interest because it benefits from the fact that by varying its shape parameter, it is possible to emulate other distributions that have been used to model ultrasound RF data and its ability to distinguish between scatterers of different concentrations and arrangements within the tissue [7]. In this work, we used it for the detection of the choroid plexus, a glomerular mass comprised of connective tissue.

The Nakagami distribution, $N(x|\mu,\omega)$, is one of the exponential probability distributions and it can be defined by two parameters: the shape (μ) and scale (ω) parameters.

$$N(x|\mu,\omega) = \frac{2\mu^{\mu}}{\Gamma(\mu)\omega^{\mu}} x^{2\mu-1} exp(\frac{-\mu}{\omega}x^2), \forall x \in \mathbb{R}_{\geq 0} \quad (1)$$

The parameters are approximated by determining the Nakagami maximum likelihood estimation (MLE) of a histogram of pixels intensities in a small rectangular region within the image. Figure 2 shows the MLE fits for a rectangular region within the choroid plexus and for a background region outside. The variation in the distribution parameters is indicative of the speckle statistics characteristic to the different tissues within the fetal brain, as it corresponds to the inhomogeneities of the tissues. The size of speckle is related to the sound wavelength, λ, and it has been found to occur in the range of approximately $0.1 \leq \lambda \leq 1.0mm$ [7].

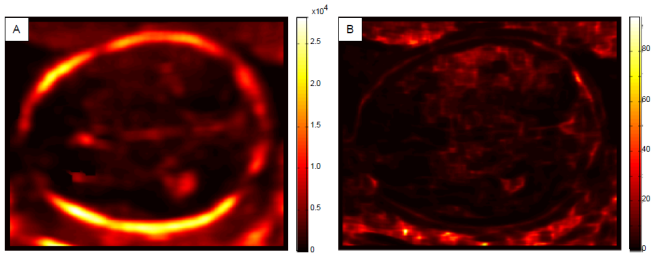


Figure 3: Nakagami μ (A) and ω (B) parameter plots for a 2D fetal brain ultrasound image.

In our images, the resolution is approximately $1.37 \times 1.37mm$, so we selected a patch of 16×16 pixels from which the distribution parameters were estimated and assigned to the center pixel of the patch, and the patch was shifted across the entire image to produce μ and ω parameter plots. Representative examples of parameter plots are shown in figure 3.

2.2 Local Feature Extraction

The objective of this study was to detect the rectangular region (subwindow) of the transthalamic plane (TP) with the highest likelihood of containing the anatomical object of interest (the choroid plexus). This subwindow is identified as the region with the maximal score determined by the classifier. In order to classify each subwindow as foreground (containing the choroid plexus) or background, we opted for the Haar-wavelet based approach for the selection of features to best discriminate between the positive (foreground, ζ) and negative (background, χ) subwindow from the image, I, such that $\zeta \cup \chi \subseteq I$ and $\zeta \cap \chi = \oslash$.

The feature score is determined by subtracting the sum of pixels in a rectangular region from the sum of pixels in an adjacent rectangular region of the same dimensions. This process was described in detail in [6].

2.3 AdaBoost Classifier

We used the AdaBoost algorithm [2] to select the features and to train the classifier, as it has proven to be effective in similar problems of object detection in fetal ultrasound images [4]. The AdaBoost algorithm linearly combines a collection of weak classifiers (i.e. the set of extracted features) to form a strong classifier, such that it boosts the performance of the weak classifiers. During training, the weight distributions of the features are modified after each iteration such that the error of classification is lower than that of the previous iteration.

The training data consisted of a set of positive and negative examples that were passed through the AdaBoost learning framework. The positive subwindow examples (containing the choroid plexus) were manually cropped from the ultrasound images and scaled to 100×100 pixels and the process was repeated for the negative samples. We obtained the training examples from three different feature sets: *a*) the intensity values of the ultrasound image pixels, I, *b*) the shape parameter, μ, and *c*) the scale parameter of the local Nakagami distribution, ω.

2.4 Data Acquisition

The training set for the choroid plexus detection consisted of 190 positive and 380 negative examples for each of the training sets. We found that the choroid plexii in the standard

transthalamic plane occupies approximately 2014±88.5 pixels which can be adequately encompassed by subwindows ranging from 75×75 to 117×117 pixels.

3 Results and Discussion

The AdaBoost training algorithm was implemented in MATLAB 7.13. The program was validated on 61 images and trained on 44 images of fetuses between 19^{+6} and 22^{+6} weeks of gestation.

3.1 Evaluation of Validation Data

The AdaBoost algorithm required the tuning of two parameters:*a*) the number of iterations that produce T weak learners (WL), and *b*) the final classifier threshold that yields the highest sensitivity and specificity. These parameters were selected based on Receiver Operating Characteristic (ROC) curves for each of the feature sets for a selected number of T weak learners. Table 1 summarizes the results from the ROC analysis.

Table 1: ROC Results for Different Number of Weak Learners (WL)

No. of WL	Intensity			Nakagami μ			Nakagami ω		
	AUC	Acc	Thresh	AUC	Acc	Thresh	AUC	Acc	Thresh
300	0.785	0.7213	0.6415	0.666	0.6393	0.6437	0.802	0.7049	0.6518
500	0.807	0.6557	0.6299	0.695	0.5738	0.6598	0.807	0.7213	0.629

A higher area under the ROC curve (AUC) yields a better object detector test producing fewer false positives, and vice-versa. Based on the results displayed in table 1, we selected 500 WL due to the fact that the highest accuracies are seen with this number of weak learners (particularly for intensity and Nakagami ω, highlighted in gray) with the respective thresholds for each of the feature sets (Table 1).

3.2 Evaluation of Independent Testing Data

The testing dataset consisted of 44 2D fetal ultrasound scans of gestational ages ranging between 20^{+0} and 22^{+6} weeks. The choroid plexus detection was considered to be positive when it occupied more than 50% of the automatically found subwindow by visual inspection. The performance of the three features sets (i.e. Intensity, Nakagami μ, Nakagami ω) was compared using ROC curves shown in figure 4, and the statistical results are summarized in table 2.

Table 2: ROC Results for Different Number of Weak Learners (WL)

Feature Set	AUC	Balanced Accuracy (%)
Intensity	0.889	61.36
Nakagami μ	0.678	65.91
Nakagami ω	0.862	72.73

Based on the findings, it is evident that the proposed Nakagami ω feature was able to outperform the intensity and Nakagami μ feature sets in the detection of the choroid plexus

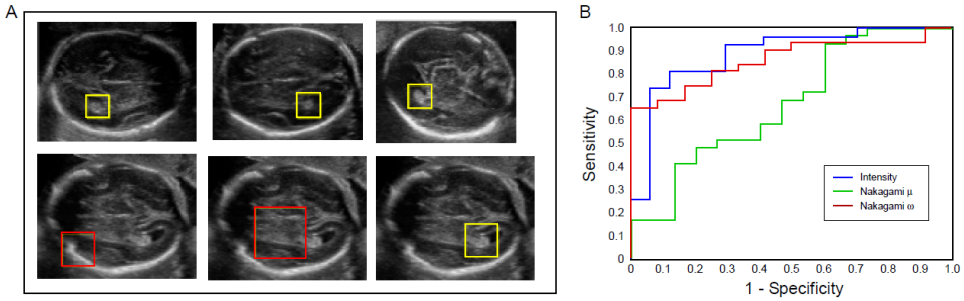


Figure 4: Choroid plexus detection results. The positively detected choroid plexuses are bounded by a yellow box, and the false positives are bounded by a red box (A). The ROC curves for 500WL are also shown (B).

with an increase in accuracy of 11.37% over the intensity feature and 6.82% over the Nakagami μ feature. This is possibly due to the fact that the scale (ω) parameter captures the characteristic spread of the speckle noise within the choroid plexus, which encapsulates textural properties as opposed to the energy of the pixel intensities captures by the shape (μ) parameter.

4 Conclusion

This paper presents a novel method for the detection of the choroid plexus in the standard transthalamic scan of the fetal head. We have tested three different feature sets on the AdaBoost object detection framework and found that the statistical model approximation yielded the best results in detection accuracy. This result can be extended to other ultrasound object detection applications, and for quality control of plane selection during the obstetrical exam.

References

[1] K.M. Dziegielewska and et al. Development of the choroid plexus. *Microsc Res Tech*, 52(1):5–20, Jan 2001.

[2] Y. Freund and R. Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In *Computational Learning Theory*. Springer, 1995.

[3] ISUOG. Sonographic examination of the fetal central nervous system. *UOG*, 29:109–116, 2007.

[4] B. Rahmatullah and et al. Quality control of fetal ultrasound images: Detection of abdomen anatomical landmarks using adaboost. In *ISBI*, pages 6–9, 2011.

[5] S.R. Turner and et al. Sonography of fetal choroid plexus cysts: detection depends on cyst size and gestational age. *J Ultrasound Med*, 22(11):1219–1227, Nov 2003.

[6] P. Viola and M.J. Jones. Robust real-time face detection. *IJCV*, 57:137–154, 2004.

[7] C. Wachinger and et al. Locally adaptive nakagami-based ultrasound similarity measures. *Ultrasonics*, 52(4):547–554, Apr 2012.

Classification of Microcalcification Clusters Using Topological Structure Features

Zhili Chen[1,5]
zzc09@aber.ac.uk

Arnau Oliver[2]
aoliver@eia.udg.edu

Erika Denton[3]
erika.denton@nnuh.nhs.uk

Caroline Boggis[4]
caroline.boggis@manchester.ac.uk

Reyer Zwiggelaar[1]
rrz@aber.ac.uk

[1]Department of Computer Science
Aberystwyth University, Aberystwyth, UK

[2]Department of Computer Architecture and Technology, University of Girona, Girona, Spain

[3]Department of Breast Imaging, Norfolk and Norwich University Hospital, Norwich, UK

[4]School of Medicine, The University of Manchester, Manchester, UK

[5]Faculty of Information and Control Engineering
Shenyang Jianzhu University, Shenyang, China

Abstract

The presence of microcalcification clusters is a primary sign of breast cancer. It is difficult and time consuming for radiologists to diagnose microcalcifications. In this paper, we present a novel method for the classification of malignant and benign microcalcification clusters in mammograms. We analyse the topology of individual microcalcifications within a cluster using multiscale morphology. A microcalcification graph is constructed to represent the topological structure of the cluster and two properties associated with the connectivity are investigated. A multiscale topological feature vector is generated from a set of microcalcification graphs for classification. The validity of the proposed method is evaluated based on the MIAS database. Using a k-nearest neighbour classifier, a classification accuracy of 95% is achieved for both manual annotations and automatic detection results. The obtained area under the ROC curve is 0.93 and 0.92 for the manual and automatic segmentation, respectively.

1 Introduction

The presence of microcalcification clusters is a primary sign of breast cancer. The radiological definition of microcalcification clusters is that at least three microcalcifications are present within a 1 cm^2 region [7, 12]. Due to its high spatial resolution, mammography enables the detection of microcalcifications at an early stage, however, it is difficult and time consuming for radiologists to distinguish malignant from benign cases. This results in a high rate of unnecessary biopsy examinations [4, 12].

In order to improve the performance of radiologists and reduce the false positive rate, computer-aided diagnosis (CAD) systems have been applied [4]. A variety of approaches have been developed for the characterisation and classification of microcalcifications. Shen *et al.* [10] and Ma *et al.* [7] developed a set of shape features to quantitatively measure the

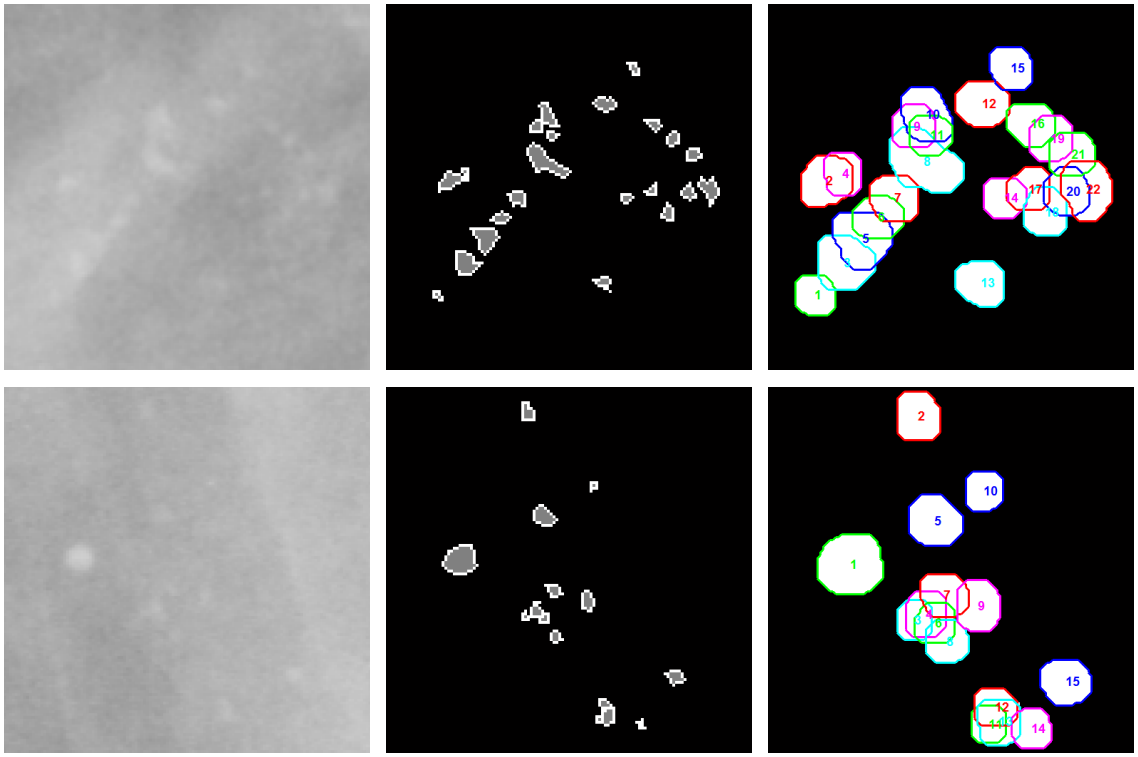


Figure 1: Example microcalcification clusters: malignant (top row) and benign (bottom row). Left column: original image patches; middle column: manually segmented microcalcifications; right column: dilated microcalcifications (at scale 7). Note that microcalcification No. 16 is not displayed in the bottom images as this falls outside the zoomed region.

roughness of individual microcalcifications. Betal *et al*. [1] analysed four shape properties of individual microcalcifications. Chan *et al*. [3] used morphological features to describe the size, shape, and contrast of individual microcalcifications and their variations within the cluster. Cluster features were used in [1, 5, 9] to describe the global properties of the cluster. The greylevel variations and texture features within the region of interest were investigated in [2] and [5]. The performance of different types of features for microcalcification classification was compared in [12], where the multiscale representation based on multiwavelet transform, shape and texture features were investigated.

Malignant microcalcifications tend to be small, numerous and densely distributed, while benign microcalcifications are generally larger, smaller in number and more diffusely distributed [11]. This difference results in various topological structures of microcalcification clusters. The distribution of microcalcifications associated with a malignant process may be different from that associated with a benign process [6]. We investigate the connectivity between individual microcalcifications using morphological operations at multiple scales.

2 Data and Method

The data used in this paper are twenty image patches (512×512 pixels) taken from the Mammographic Image Analysis Society (MIAS) database [13], each containing a microcalcification cluster. There are nine malignant and eleven benign cases (biopsy proven) in the dataset. All individual microcalcifications have been manually segmented by an expert radiologist. The median number of microcalcifications in the clusters is 27. There are a

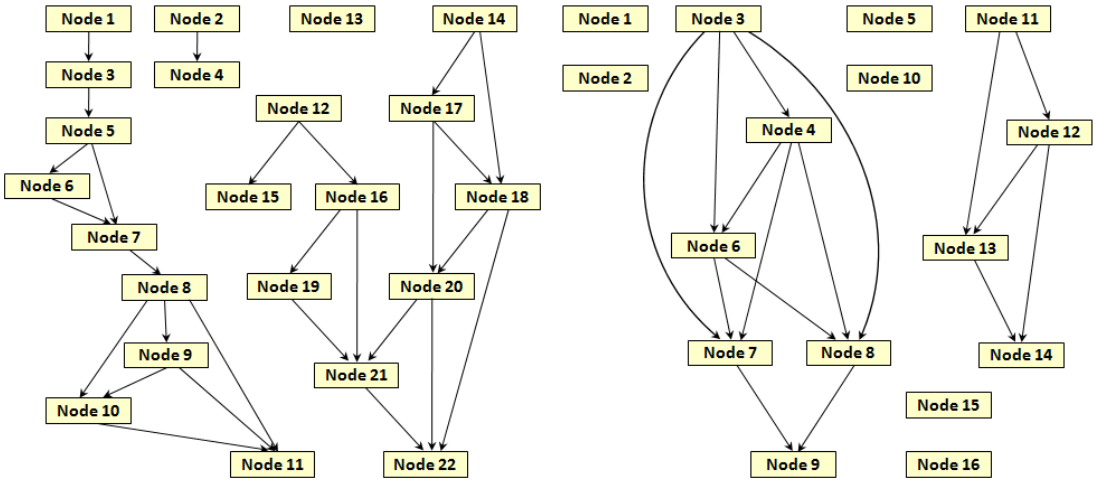


Figure 2: Microcalcification graphs: (a) malignant ($n_s = 4$, $\delta_s = 1.23$); (b) benign ($n_s = 8$, $\delta_s = 1.00$). The numbering of nodes is consistent with the sequence number in Figure 1.

few outliers and 80% of the clusters are within the 6 to 62 range. Figure 1 shows example microcalcification clusters and corresponding manual annotations. In addition, an approach developed by Oliver *et al*. [8] is used to detect microcalcifications. Probability images are obtained where high values indicate microcalcifications. To segment microcalcifications, the probability images are binarised using a threshold determined using ROC analysis.

2.1 Morphological Operation

Multiscale morphological dilation is performed on individual microcalcifications using a disk-shaped structuring element with a radius equal to the scale (see the right column of Figure 1, where each individual microcalcification is ordered with a sequence number and the boundaries of dilated microcalcifications are displayed using different colours). It indicates that the morphological dilation adds neighbouring pixels to the boundaries of individual microcalcifications, resulting in changes in the connectivity between individual microcalcifications within clusters.

2.2 Microcalcification Graph

The topology of individual microcalcifications within the cluster is represented in graphical form. A microcalcification graph is constructed based on the spatial connectivity relationship between microcalcifications, where each node represents one microcalcification, and two nodes are linked by an edge if the corresponding two microcalcifications are connected or overlap with each other. Here, we generate a directed graph where the nodes are ordered according to the spatial location of the corresponding microcalcifications in the image patch, and two connected nodes are linked by a directed edge from the smaller to the larger numbered node. The resulting graphs of dilated microcalcifications in Figure 1 are shown in Figure 2. We focus on two properties of the microcalcification graph. The first property is the number of independent subgraphs within the graph, which represents the number of independent connected components within the microcalcification cluster. The second property is the degree of each node defined as the number of edges starting from the node, which describes the connectivity of the corresponding microcalcification with its neighbours.

We define an upper-triangular adjacency matrix to encode the directed microcalcification graph, denoted by $A = (a_{ij})$, $a_{ij} \in \{0,1\}$, $i,j = 1,\ldots,m$, where m is the number of nodes within the graph. Node i and node j are connected if $a_{ij} = 1$. Node i is the source node and node j is the sink node. A source node i is called a *root* node if $\sum_{k=1}^{m} a_{ki} = 0$. A sink node j is called a *terminal* node if $\sum_{k=1}^{m} a_{jk} = 0$. A *path* from node i to node j is defined as a sequence of nodes starting from node i and ending with node j. The number of independent subgraphs (denoted by n) is determined by traversing the graph. We traverse the graph starting at each *root* node and explore as far as possible along each *path* until arriving at the *terminal* node. The traversal sequences including common nodes are combined into a single sequence. The number of the final sequences is the number of independent subgraphs. The degree of node i (denoted by $\delta(i)$) is computed by $\delta(i) = \sum_{k=1}^{m} a_{ik}$.

2.3 Multiscale Topological Feature Vector

We generate a set of microcalcification graphs $G = (G_0, G_1, \ldots, G_{S-1})$ based on dilated microcalcifications at S scales. We extract properties from G and form two feature vectors $N = (n_0, n_1, \ldots, n_{S-1})$ and $\Delta = (\delta_1, \delta_2, \ldots, \delta_{S-1})$, where $n_s (s = 0,1,\ldots,S-1)$ is the number of independent subgraphs at scale s, and $\delta_s (s = 0,1,\ldots,S-1)$ is the average degree of nodes at scale s, computed by $\frac{1}{m}\sum_{i=1}^{m} \delta(i)_s$. We normalise N and Δ by n_s/m and $\delta_s / \max \delta(i)_s$, where $\max \delta(i)_s$ is the maximum degree of nodes at scale s. Finally, the two normalised vectors are concatenated into a single feature vector, which can be used for the classification of malignant and benign microcalcification clusters.

3 Experimental Results

For each microcalcification cluster, we analysed the morphological topology of microcalcifications at 129 scales ($s = 0,1,\ldots,128, S = 129$). As described in Section 2, we computed n_s and δ_s using the corresponding microcalcification graph at scale s. The dimensionality of the obtained multiscale topological feature vectors was 258.

A k-nearest neighbour (kNN) classifier and a leave-one-out methodology were used for classification. The Euclidean distance was used to measure the similarity between feature vectors. The sequential forward selection (SFS) algorithm was applied to select the most discriminating features. In the feature selection process, SFS was performed based on the training set excluding the testing sample to avoid bias. For different training sets, on average approximate 90 features were retained for the feature set (manual data) extracted from the manual annotations and approximate 120 features were retained for the feature set (CAD data) extracted from the automatic detections. The most frequently selected features (occurrence > 80%) are mainly within the scale range of $[1,32]$ for the manual data and $[28,68]$ for the CAD data. For the manual data, using the unreduced feature space, the best classification accuracy (CA) was 90% for $k = 3$, where nine of the eleven benign cases were classified correctly without misclassifying any malignant cases. Using the reduced feature space, the best CA was increased to 95% with one benign case misclassified. For the CAD data, using the unreduced feature space, the best CA was 85% for $k = 5$ with two benign and one malignant cases misclassified. Using the reduced feature space, the best CA was 95% for $k = 3$ with one malignant case misclassified. This indicates the method is robust with regard to microcalcification segmentation and performs well on the CAD detection results.

Table 1: Comparison of our achieved results with those obtained by some related work.

Method	Database	# Case	Feature	Classifier	Result
[10]	unknown	18	shape	kNN	$CA = 100\%$
[1]	Liverpool	38	shape/cluster	kNN	$A_z = 0.79 A_z = 0.84$
[7]	DDSM	183	shape	α_{max}	$A_z = 0.96$
[3]	unknown	145	morphological	LDC	$A_z = 0.79$
[5]	unknown	191	texture & cluster	ANN	$A_z = 0.86$
[9]	MIAS	25	cluster	SVM	$A_z = 0.81$
[2]	unknown	54	texture	ANN	$A_z = 0.88$
[12]	Nijmegen	103	multiwavelet	kNN	$A_z = 0.89$
Our	MIAS (manual)	20	multiscale topology	kNN	$CA = 95\%$ $A_z = 0.93$
	MIAS (CAD)	20	multiscale topology	kNN	$CA = 95\%$ $A_z = 0.92$

To quantitatively assess the classification performance of the multiscale topological features, a ROC curve was constructed. The ROC curve represents the trade-off between the true positive rate (TPR) against the false positive rate (FPR). The construction of the ROC curve is based on a decision criterion which can be regarded as a threshold to decide a test sample as either positive or negative. We defined a malignancy measure (denoted by M) as the decision criterion based on the kNN classifier. The malignancy measure M of a testing microcalcification cluster was defined to be the number of malignant clusters among its k nearest neighbours, ranging from 0 to k. Thus, a threshold L was set from -1 to k, and the testing cluster was classified as malignant if M was larger than L. When $L = -1$, all the microcalcification clusters were classified as malignant with TPR and FPR equal to 1. At the other extreme, when $L = k$, all the microcalcification clusters were classified as benign with TPR and FPR equal to 0. The remaining TPR and FPR were obtained by varying L from 0 to $k-1$. This produced $k+2$ points of TPR and FPR. Finally, the area under the ROC curve (denoted by A_z) was computed using the trapezoidal rule. We tested three k values: $k = 3$, $k = 5$, and $k = 7$. For the manual data, the obtained A_z was 0.88, 0.91, 0.85, and 0.91, 0.93, 0.87, using the unreduced and reduced feature space, respectively. For the CAD data, the obtained A_z was 0.81, 0.88, 0.80, and 0.88, 0.92, 0.86, using the unreduced and reduced feature space, respectively.

We compared our proposed method with some related publications. Table 1 shows a summary of the comparison. It is shown that our method provides comparative results for both manual and CAD segmentation. Note that the various approaches use different images taken from different databases, and therefore it is a qualitative comparison.

4 Conclusions

This is a novel approach to investigate the morphological topology and connectivity of microcalcifications for discriminating malignant from benign clusters. Unlike most features in previous publications extracted at a single scale, a representation covering the multiscale characteristics was developed in this paper. To evaluate the validity of this method, we used manually segmented microcalcifications and CAD detection results. Good classification results were obtained for both types of data. This indicates the robustness of this method to detection errors and the potential application in CAD systems. As feature work, alternative

approaches to feature selection (e.g. genetic algorithm) and other classifiers (e.g. decision tree, artificial neural network, and support vector machine) will be employed for classification. The definition of a similarity measure between graphs will be further investigated in order to realise classification using the graph based representation directly without generating feature vectors. In addition, further evaluation using a larger dataset taken from the DDSM database and a dataset of full-field digital mammograms is ongoing.

References

[1] D. Betal et al. Segmentation and numerical analysis of microcalcifications on mammograms using mathematical morphology. *British Journal of Radiology*, 70(837):903–917, 1997.

[2] H. P. Chan et al. Computerized classification of malignant and benign microcalcifications on mammograms: Texture analysis using an artificial neural network. *Physics in Medicine and Biology*, 42:549–567, 1997.

[3] H. P. Chan et al. Computerized analysis of mammographic microcalcifications in morphological and texture feature spaces. *Medical Physics*, 25(10):2007–2019, 1998.

[4] H. D. Cheng et al. Computer-aided detection and classification of microcalcifications in mammograms: A survey. *Pattern Recognition*, 36(12):2967–2991, 2003.

[5] A. P. Dhawan et al. Analysis of mammographic microcalcifications using gray-level image structure features. *IEEE TMI*, 15(3):246–259, 1996.

[6] S. A. Feig et al. Evaluation of breast microcalcifications by means of optically magnified tissue specimen radiographs. *Recent Results in Cancer Research*, 105:111–123, 1987.

[7] Y. Ma et al. A novel shape feature to classify microcalcifications. In *Proc. ICIP*, pages 2265–2268, 2010.

[8] A. Oliver et al. Automatic microcalcification and cluster detection for digital and digitised mammograms. *Knowledge-Based Systems*, 28:68–75, 2012.

[9] A. Papadopoulos et al. Characterization of clustered microcalcifications in digitized mammograms using neural networks and support vector machines. *Artificial Intelligence in Medicine*, 34(2):141–150, 2005.

[10] L. Shen et al. Application of shape analysis to mammographic calcifications. *IEEE TMI*, 13(2):263–274, 1994.

[11] E. A. Sickles. Breast calcifications: Mammographic evaluation. *Radiology*, 160(2): 289–293, 1986.

[12] H. Soltanian-Zadeh et al. Comparison of multiwavelet, wavelet, Haralick, and shape features for microcalcification classification in mammograms. *Pattern Recognition*, 37 (10):1973–1986, 2004.

[13] J. Suckling et al. The Mammographic Image Analysis Society digital mammogram database. In *International Congress Series 1069*, pages 375–378, 1994.

A Statistical Approach to Modelling Breast Tissue Appearance in Mammograms

Zhili Chen[1,3]
zzc09@aber.ac.uk

Erika Denton[2]
Erika.denton@nnuh.nhs.uk

Reyer Zwiggelaar[1]
rrz@aber.ac.uk

[1]Department of Computer Science
Aberystwyth University, Aberystwyth, UK

[2]Department of Breast Imaging, Norfolk and
Norwich University Hospital, Norwich, UK

[3]Faculty of Information and Control Engineering
Shenyang Jianzhu University, Shenyang, China

Abstract

In this paper, we present an approach to build models of breast tissue appearance in mammograms. Mammographic tissue is modelled based on a statistical analysis of local appearance. We investigate five strategies by using different types of local features, covering aspects of intensity, texture, and geometry. A visual dictionary is generated to summarise local tissue appearance with descriptive "words". The global appearance of the breast is represented as an occurrence histogram over the dictionary. The resulting histogram models can be applied to breast density classification. The validity is qualitatively and quantitatively evaluated using the full MIAS database. The consensus of three experts according to the BIRADS criterion is used as the classification ground truth. We test the performance of each individual strategy and the combination of all strategies. The results indicate that our approach has potential for mammographic risk assessment.

1 Introduction

Many studies have indicated that there is a strong correlation between breast tissue density/patterns and breast cancer risk. Therefore, modelling mammographic tissue appearance is beneficial for the qualitative perception of breast tissue patterns, quantitative analysis of breast density, and automated mammographic risk assessment. Recently, a variety of approaches have been developed to characterise breast tissue in mammograms [4, 7, 8, 12]. In general, intensity information is used to describe tissue density and texture information is used to represent tissue patterns. In this paper, we present a statistical approach to model breast tissue appearance over the whole breast. We focus on local tissue appearance in terms of intensity, texture, and geometry. Local features are extracted from the neighbourhood of breast tissue pixels and are statiscally analysed to build overall models of breast tissue.

2 Methodology

The schema of our method is shown in Fig. 1. Firstly, as a preprocessing step, the breast region is segmented using the approach in [1]. Subsequently, features are extracted from the

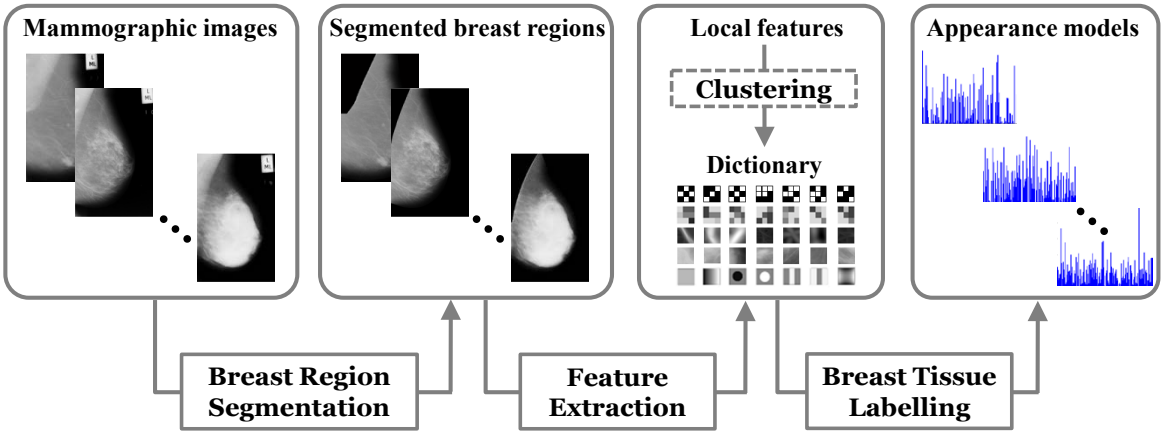


Figure 1: The schema of modelling breast tissue appearance based on the statistical analysis.

local neighbourhood of breast tissue pixels within the breast region. After that, a visual dictionary of local tissue appearance is generated in two ways: one is aggregating an exhaustive configuration of a certain type of local features; the other is performing an initial training step to "learn" clusters of local features. Finally, each breast tissue pixel is labelled by searching for the nearest "word" in the dictionary. Each mammographic image is represented as an occurrence histogram of the "words" in the dictionary. We investigate five different strategies for local features, including local binary patterns, local greylevel appearences, local geometric structures, joint filter responses, and raw image patches, respectively.

Local Binary Patterns (LBP) were first proposed in [6]. Local appearance of breast tissue is encoded into a set of binary values. The greylevel value of the centre pixel is substracted from the local neighbourhood, and a binary label is assigned to each neighbouring pixel according to the difference sign. The resulting binary pattern is transformed into a unique LBP number by $LBP = 1 + \sum_{p=0}^{P-1} s(g_p - g_c)2^p$, where g_c is the greylevel value of the centre pixel, g_p denotes the greylevel value of the p^{th} pixel in the local neighbourhood, and $s(x) = 1$ if $x \geq 0$ else $s(x) = 0$. Thus, each LBP number corresponds to a unique local binary pattern, and all possible patterns comprise the visual dictionary of local appearance. An LBP histogram is populated by counting the occurrences of LBP numbers at every pixel.

The Local Greylevel Appearance (LGA) based approach was presented in [12]. Breast tissue appearance is modelled by analysing the joint greylevel distribution of the local neighbourhood. The local greylevel appearance is transformed into a unique LGA number by $LGA = 1 + \sum_{i,j} N_g{}^{counter(i,j)} I(i,j)$, where N_g is the greylevel resolution, $counter(i,j)$ is the sequence number of pixel (i,j) within the neighbourhood, and $I(i,j)$ is the greylevel value of pixel (i,j). Thus, each LGA number corresponds to a unique local greylevel appearance, and all possible greylevel appearances comprise the visual dictionary. An LGA histogram containing the combination of LGA numbers and corresponding occurrences is generated.

Basic Image Features (BIF) were defined in [2]. A second-order family of six Gaussian derivative filters are used to analyse local appearance of breast tissue with respect to multi-scale geometric structures. Mammographic images are convolved with the Gaussian filter bank at multiple scales. Seven BIFs are defined, each corresponding to a distinct type of local geometric structures. For each breast tissue pixel, the corresponding geometric structure is determined according to the largest BIF computed with its local neighbourhood. The configuration of the seven geometric structures across multiple scales is encoded into a unique BIF column, and all possible configurations comprise the visual dictionary. Breast tissue

appearance is modelled by counting the occurrences of BIF columns over the whole breast.

In this paper, we use two types of textons. The first type of textons are generated by filtering a set of training mammograms with a filter bank and clustering the joint filter responses. The cluster centres are considered as textons. We use the MR8 filter bank defined in [10], where only eight filter responses are retained. The aggregated filter responses of breast tissue pixels over all the training mammograms are clustered using the classic K-Means algorithm. The resulting textons comprise the dictionary of local tissue appearance. Breast tissue appearance is modelled by a frequency histogram of textons. The second type of textons are generated by clustering raw image patches extracted from the training mammograms [11]. The procedure of learning textons is similar as described above. A frequency histogram of textons based on source image patches constructs the appearance model of breast tissue. It has been demonstrated in [11] that the image-patch based textons can provide superior performance to the filter-response based textons for the sake of texture classification. We refer to the two modelling strategies based on the MR8 and image-patch textons as Texton I and Texton II, respectively. Note that the preprocessing step in [10, 11] is not performed in Texton I and Texton II, and the filter responses in Texton I are not normalised by Weber's law as in [10, 11], in order to retain the original intensity correlation between mammograms.

3 Experiments and Results

In our experiments, the Mammographic Image Analysis Society (MIAS) database [9] was used, which contains 322 Medio-Lateral Oblique (MLO) mammograms taken from 161 women. The original spatial resolution is $50\mu m \times 50\mu m$ per pixel. Due to memory and efficiency reasons, we downsampled the full resolution to $800\mu m \times 800\mu m$ per pixel. Three experts classified 321 available mammograms (mdb295ll was excluded for historical reasons) into four Breast Imaging Reporting and Data System (BIRADS) density categories [5]. The consensus between individual classification decisions was considered as the ground truth. We built appearance models for each mammogram using the five strategies described above.

When modelling breast tissue appearance, to eliminate bias caused by mammogram edges and the breast-background boundary, we only focused on the pixels with neighbourhoods entirely located within the breast region, which can be identified automatically based on the pre-segmented breast region. For LBP, we used a 3×3 neighbourhood. For LGA, we also used a 3×3 neighbourhood and reduced the greylevel resolution to 16. For BIF, we convolved mammographic images with the Gaussian filter bank at four scales. For the two texton based methods, at the texton learning stage, we randomly selected 40 mammograms from the database as the training set. For texton II, we extracted 3×3 image patches and rearranged the pixels in row order for the K-means clustering. For each individual method, we learned 160 textons from the training set. Thus, the number of "words" generated for each type of local feature was 2^9, 16^9, 7^4, 160 and 160, respectively

For qualitative evaluation, we display the results in the form of label maps. To generate a label map, at each breast tissue pixel, we searched for the nearest "word" to the extracted local feature across the dictionary, and labelled each pixel according to the response "word". Fig. 2 shows the resulting label maps of example mammograms covering the four BIRADS categories. For LBP, LGA and BIF, the "words" were sorted according to the order of LBP number, LGA number and BIF column. For texton I and texton II, the "words" were sorted according to the magnitude value of textons in ascending order before labelling the pixels. It is shown that LBP is sensitive to noise and small textures, and there are no homogeneous

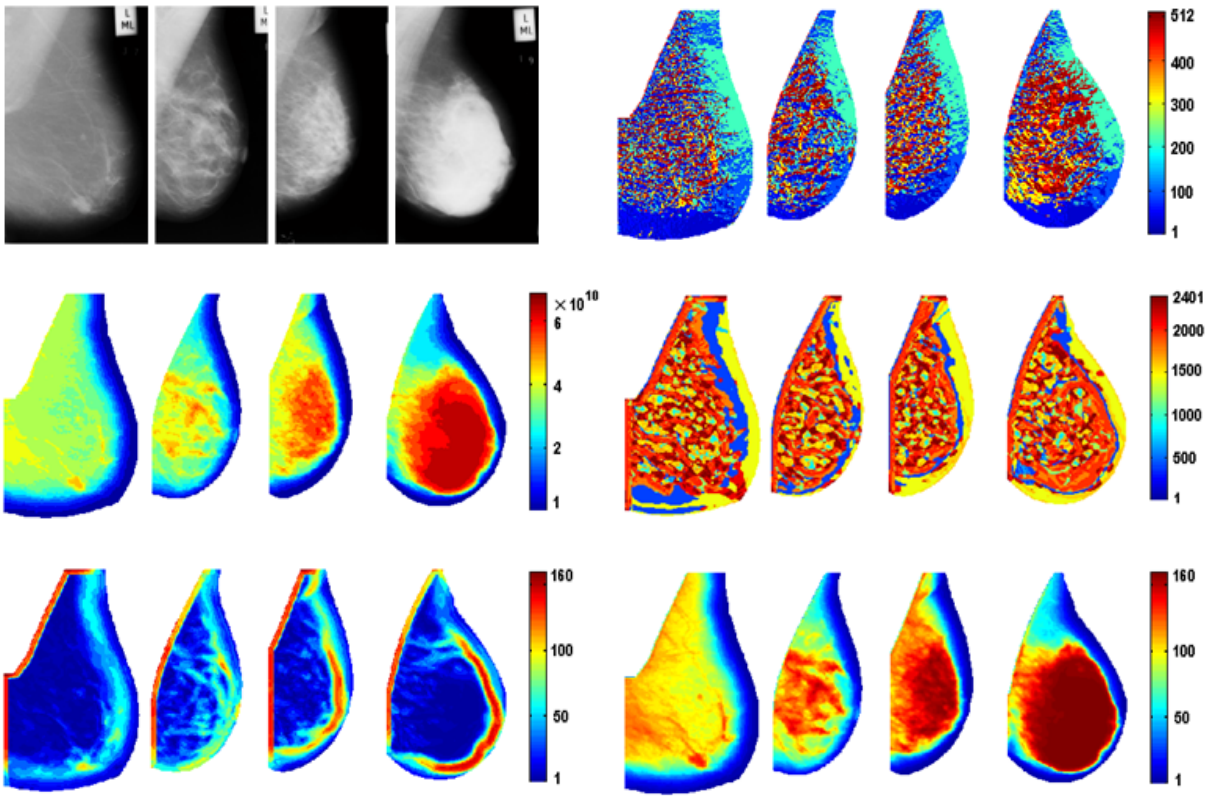


Figure 2: Example mammograms and resulting label maps using different strategies: top row: original mammograms and LBP; middle row: LGA and BIF; bottom row: Texton I and Texton II. For each batch, from left to right corresponds to an increasing BIRADS category.

regions obtained within the breast area. Both LGA and Texton II indicate realistic segmentation with respect to tissue density. Texton I has strong responses to the boundary between dense and fatty tissue regions, and relatively homogeneous regions are obtained within the dense/fatty tissue. For BIF, cell-like regions are obtained within the breast area, corresponding to different types of geometric structures.

For quantitative evaluation, we applied the resulting models in the form of occurrence histograms (L_1 normalised) to breast density classification. As described in Section 2, the occurrence histograms were populated by spanning the full dictionary. This could result in sparse histograms in the real observation of breast tissue. Therefore, we avoided using the full range of the histograms for the classification. We sorted the histogram bins in descending order to select the dominant "words" from the dictionary. We chose the most frequently used "words" occupying 99% occurrences as the dominant "words". Thus, the histogram bins corresponding to the "words" never or rarely referenced were removed. In our experiments, it was indicated that the feature space dimensionality was significantly reduced after compressing the visual dictionary, especially for LBP, LGA, and BIF.

A leave-one-woman-out methodology was used for the evaluation. We applied the multi-class boosting algorithm for the classification. We chose to use the Gentleboost algorithm [3] to train a classifier for each strategy, where the weak learners used in each of serial rounds are decision stumps which are regarded as degenerate decision trees with a single node. The created classifiers can provide continuous-valued outputs, which are interpreted as posterior probabilities. For a testing mammogram with a histogram-like model x, the classifier output corresponding to the class B_j is the probability of x to belong to B_j, denoted by $p_{B_j}(x)$. To

Table 1: Confusion matrices for breast density classification using different strategies.

LBP

	BIRADS	I	II	III	IV	CA
Truth	I	72	13	2	0	83%
	II	14	60	27	2	58%
	III	7	13	66	8	70%
	IV	0	5	21	11	30%

LGA

	BIRADS	I	II	III	IV	CA
Truth	I	69	15	3	0	79%
	II	19	47	37	0	46%
	III	1	25	64	4	68%
	IV	1	4	5	27	73%

BIF

	BIRADS	I	II	III	IV	CA
Truth	I	74	13	0	0	85%
	II	14	63	25	1	61%
	III	1	27	54	12	57%
	IV	0	2	19	16	43%

Texton I

	BIRADS	I	II	III	IV	CA
Truth	I	72	15	0	0	83%
	II	12	73	18	0	71%
	III	2	26	55	11	59%
	IV	0	4	11	22	59%

Texton II

	BIRADS	I	II	III	IV	CA
Truth	I	76	10	1	0	87%
	II	11	71	18	3	69%
	III	1	23	64	6	68%
	IV	1	2	14	20	54%

Comb.

	BIRADS	I	II	III	IV	CA
Truth	I	79	8	0	0	91%
	II	11	78	14	0	76%
	III	1	19	70	4	74%
	IV	1	1	11	24	65%

make the sum of the outputs over all the four BIRADS classes equal to 1, they are normalised by $P(B_j \mid x) = e^{p_{B_j}(x)} / \sum_{c=1}^{4} e^{p_{B_c}(x)}$, where $P(B_j \mid x)$ is the normalised value of $p_{B_j}(x)$.

Table 1 shows the classification results obtained by using the five modelling strategies. The overall classification accuracy (CA) is 65%, 64%, 64%, 69%, and 72%, for LBP, LGA, BIF, Texton I, and Texton II, respectively. The best classification result is obtained by Texton II, which is followed by Texton I, providing the second-best result, while LGA and BIF perform worst among these five methods. In addition, we investigated the performance of combining the outputs of the five individual classifiers. A weighted average combination rule was used to compute the total probability for each class. For a testing mammogram, the total probability corresponding to the class B_j (denoted by $P_{sum}(B_j)$) is obtained by $P_{sum}(B_j) = \sum_{t=1}^{5} w_t P(B_j \mid x_t)$, where $P(B_j \mid x_t)$ is the output of the t^{th} classifier, and w_t is the corresponding weight value. We set $w_1 = 0.16$, $w_2 = 0.12$, $w_3 = 0.12$, $w_4 = 0.32$, and $w_5 = 0.28$ experimentally, but small variations provided similar results. The obtained overall CA is 78%, which indicates better performance compared with those obtained by the individual classifiers. The resulting confusion matrix is also shown in Table 1.

We compared the obtained results with publications where the BIRADS criterion was used for breast density classification. Petroudi *et al.* [8] modelled parenchymal patterns of the whole breast with a statistical distribution of textons. They obtained an overall CA of 76% for the Oxford Database. Oliver *et al.* [7] extracted morphological and texture features from dense and fatty tissue regions. They obtained an overall CA of 77% for the MIAS database, which increased to 86% when the Bayesian combination of the kNN classifier and the C4.5 decision tree was used. He *et al.* [4] developed a number of mammographic image segmentation methods for mammographic risk assessment. The best classification accuracy that they recently obtained was 75%. It is shown that our results are comparable to related publications. Note that the same database and the same classification ground truth were used in [4, 7], which enables a direct comparison.

4 Discussion and Conclusions

There are several parameters in the five modelling strategies. One common parameter in LBP, LGA, and Texton II is the size of the local neighbourhood of breast tissue pixels.

Using a large neighbourhood will drastically increase the feature space dimensionality. In LBP and LGA, the number of histogram bins grows exponentially as the neighbourhood size increases, which would raise the risk of overfitting the data and result in sparser histograms. A second parameter in LGA is the greylevel resolution, which also has a significant effect on the feature space dimensionality. For the texton based strategies, increasing the number of textons would also result in an increased risk of overfitting for the K-Means clustering. The scale parameters in the filter banks and the image spatial resolution involve the multi-resolution aspect, which could be further investigated.

In summary, we have presented an approach to modelling breast tissue based on a statistical analysis of local appearance. We generated a visual dictionary of generic breast tissue appearance by aggregating local features across different density classes. We investigated five types of local features and evaluated their performance quanlitatively and quantitatively. To our knowledge, this work is the first attempt to combine different modelling strategies for breast density classification. The experimental results indicate the validity of the approach.

References

[1] Z. Chen and R. Zwiggelaar. Segmentation of the breast region with pectoral muscle removal in mammograms. In *Proc. MIUA*, 2010.

[2] M. Crosier and L. D. Griffin. Using basic image features for texture classification. *IJCV*, 88(3):447–460, 2010.

[3] J. Friedman et al. Additive logistic regression: a statistical view of boosting. *The Annals of Statistics*, 28(2):337–407, 2000.

[4] W. He et al. Mammographic image segmentation and risk classification using a novel texture signature based methodology. In *LNCS*, volume 6136, pages 526–533, 2010.

[5] American College of Radiology. *Illustrated Breast Imaging Reporting and Data System BIRADS*. Amer. College of Radiol, 1998.

[6] T. Ojala et al. Multiresolution grey-scale and rotation invariant texture classification with local binary patterns. *PAMI*, 24(7):971–987, 2002.

[7] A. Oliver et al. A novel breast tissue density classification methodology. *TITB*, 12(1): 55–65, 2008.

[8] S. Petroudi et al. Automatic classification of mammographic parenchymal patterns: A statistical approach. In *Proc. EMBC*, pages 798–801, 2003.

[9] J. Suckling, *et al*. The Mammographic Image Analysis Society digital mammogram database. In *International Congress Series 1069*, pages 375–378, 1994.

[10] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *IJCV*, 62(1):61–81, 2005.

[11] M. Varma and A. Zisserman. A statistical approach to material classification using image patch exemplars. *PAMI*, 31(11):2032–2047, 2009.

[12] R. Zwiggelaar. Local greylevel appearance histogram based texture segmentation. In *LNCS*, volume 6136, pages 175–182, 2010.

Efficient automatic segmentation of vessels

Adriana Romero[1]
aromero@cvc.uab.es
Simeon Petkov[2]
spetkov@cvc.uab.es
Carlo Gatta[1]
carlo.gatta@ub.edu
Dr. Manel Sabaté[3]
masabate@clinic.ub.es
Petia Radeva[1]
petia.ivanova@ub.edu

[1] Universitat de Barcelona
Barcelona
Spain

[2] Centre de Visió per Computador
Bellaterra
Spain

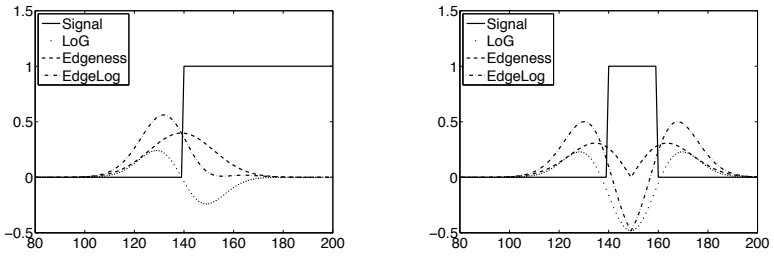
[3] Institut Clinic del Torax
Hospital Clinic Barcelona
Spain

Abstract

The segmentation of tubular structures is still an open field of investigation, particularly in medical imaging, where the quality of the image is poor with respect to natural images. Despite the quality of state-of-the-art segmentation methods, little effort has been devoted to the computational efficiency of the algorithms. Efficiency is an important topic, since intra-operative computer assisted interventions require near real-time performance. In this paper, we present a simple, yet effective, algorithm that efficiently segments vessels in 2D images. The algorithm requires no initialization and has a computational cost of $\mathcal{O}(SN\log N)$, where S is the number of scales and N is the number of image pixels. Results on the DRIVE dataset show that the proposed method has near state-of-the-art performance with very little computational burden.

1 INTRODUCTION

There is a vast literature in the area of vessel analysis and segmentation. A detailed review of existing vessel extraction techniques is given in [5]. Among the existing techniques, machine learning-based techniques may use knowledge about appearance of blood vessels in the segmentation process [11]. Hessian-based techniques analyze the eigenvalues of the Hessian matrix to perform the vessel segmentation [3]. Other techniques use medialness functions [6] as vessel detection filters. The great majority of these methods have high computational costs and/or require manual initialization of one or more seed points. We propose an efficient algorithm which is based on the following ideas: (1) detect blobs, tubular structures and edges through linear filtering, without having to compute the Hessian matrix for all pixels/voxels; (2) select a number of seeds based on the novel EdgeLoG measure and reject blob-like points by means of eigenvalues analysis, so that we limit the Hessian and eigenvalues computation only to potential seed points; (3) use a region growing approach in which edge information is used to stop the propagation. We validate our method on the



(a) (b)

Figure 1: Edgeness, Laplacian of Gaussian and EdgeLoG responses of (a) a 1-D edge signal and (b) a 1-D tubular signal.

DRIVE dataset[1], obtaining a kappa value of 0.6991, which compares well with much more sophisticated state-of-the-art methods.

2 METHOD

The method combines an edge detection operator with a blob detection operator and exploits their properties to obtain a negative signal in the presence of vessels. The approach computes the Edgeness and the Laplacian of Gaussian of an image. Edgeness obtains a positive response in the presence of edges and the Laplacian of Gaussian obtains a positive/negative response where there is contrast and 0 in constant regions. The combination of both signals is able to distinguish between tubular structures and edges and provides a negative response with a local minimum in the middle of the vessels.

Fig. 1 outlines the basic idea of the approach on a 1D toy problem. Fig. 1(a) shows the response of the operators to an edge signal. The response of the combined operator is always positive. Fig. 1(b) shows the response of the operators to a tubular-like signal. In this case, the response of the combined operators has a negative local minimum in the middle of the tubular structure. Then, we select reliable local minima in the middle of tubular structures and use them as seeds in a region growing process. Since the Laplacian of Gaussian signal can detect blobs, we reject all seeds that represent blob-like structures.

2.1 EdgeLoG computation

Let $I(x,y)$ be a grayscale image of any size. The scale space representation of $I(x,y)$ is defined as the image convolved with a family of Gaussian kernels [7]: $L^{(1)}(x,y;\sigma_i) = G(x,y;\sigma_i) * I(x,y)$ where $G(x,y;\sigma_i)$ is a Gaussian kernel of standard deviation σ_i. We also define a k-modified scale space: $L^{(k)}(x,y;k\sigma_i) = G(x,y;k\sigma_i) * I(x,y)$ where k is a constant and $G(x,y;k\sigma_i)$ is a Gaussian kernel of standard deviation $k\sigma_i$. We define the Edgeness signal as $E(x,y) = \max_{\sigma_i} ||\nabla L^{(\cdot)}(x,y;\sigma_i)||$ where $L_x^{(\cdot)}(x,y;\sigma_i)$ and $L_y^{(\cdot)}(x,y;\sigma_i)$ are the normalized first derivatives of $L^{(\cdot)}(x,y;\sigma_i)$ as in [7]. The Laplacian of Gaussian is $\nabla^2 L^{(\cdot)}(x,y;\sigma_i) = L_{xx}^{(\cdot)}(x,y;\sigma_i) + L_{yy}^{(\cdot)}(x,y;\sigma_i)$ where $L_{xx}^{(\cdot)}(x,y;\sigma_i)$ and $L_{yy}^{(\cdot)}(x,y;\sigma_i)$ are the normalized scale-space second derivatives [7].

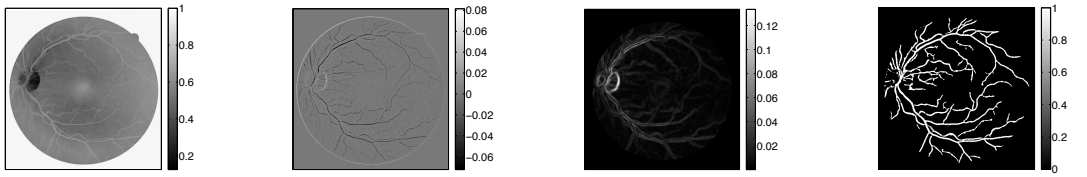
[1] http://www.isi.uu.nl/Research/Databases/DRIVE/

We propose a new measure, named EdgeLoG which conveniently combines the Laplacian of Gaussian and the Edgeness signals.

$$EL(x,y) = \min_{\sigma_i}(\nabla^2 L^{(1)}(x,y;\sigma_i) + ||\nabla L^{(k)}(x,y;\sigma_i)||) \quad (1)$$

The rationale behind this new operator is the following: (1) the Laplacian of Gaussian has a strong negative response in the presence of bright blobs, tubular structures and edges; (2) Edgeness obtains a positive response in the presence of edges and is 0 in the middle of vessels; (3) we remove the negative response of the Laplacian of Gaussian at edges while maintaining negative values in the middle of the vessels, using the Edgeness response (see e.g. Fig. 1). However, using the same standard deviation for both detectors has proven not to work properly. Therefore we introduce a constant k which modifies the Edgeness response and compensates responses to edges and tubular structures. The constant k can be computed as the minimum value that provides a negative response in ridges and a strictly positive response at edges.

To allow the detection of different vessel sizes, we compute the response of the EdgeLoG operator at multiple scales and we select the minimum EdgeLoG response over scales.



(a) Input image (b) EdgeLog response (c) Edgeness (d) Final segmentation

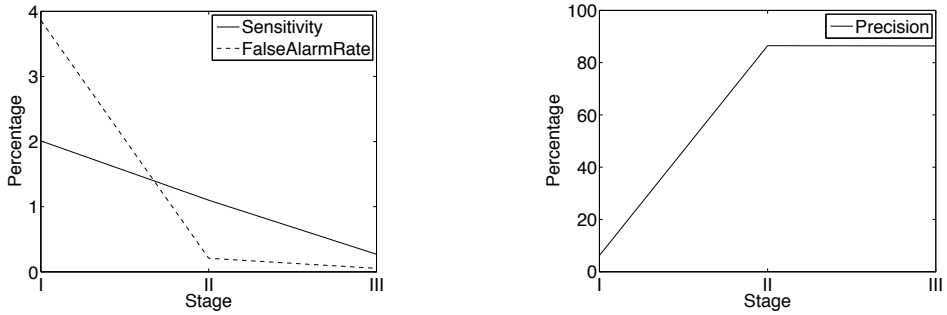
Figure 2: 2D Segmentation of retinal vessels.

2.2 Seed selection/pruning

STAGE I: Once the EdgeLoG signal is obtained, we search for local minima to define the initial set of seeds.

STAGE II: Not all local minima are reliable seeds. To reduce the number of seeds, we compute an image dependent threshold $\varepsilon < 0$ using the Otsu method [10] on the EdgeLoG intensities of the previously computed seeds, pruning seeds with EdgeLoG $> \varepsilon$. This process already rejects a consistent quantity of seeds (96.07 % on average).

STAGE III: Some of these minima may belong to blob-like structures. To identify blob-like seeds, we compute the Hessian matrix of the seed locations at the respective dominant scale, and compute their eigenvalues $|\lambda_1| \leq |\lambda_2|$. The ratio $r = |\lambda_1|/|\lambda_2| \in [0\,1]$ is a nice indicator of the blob-ness and saddle-ness of the structure. We want to reject as many non-vessel structures as possible. Therefore, we remove seed points having $r > 0.5$. In this way, we limit the possibility of using some false positives as seeds, which could induce severe errors in the region growing process. Nonetheless, removing a large amount of true positives is not a problem to our algorithm since the proposed region growing approach can produce satisfactory results with just one seed per vessel branch (see Fig. 2(d)). Fig. 3(a) and Fig. 3(b) present respectively the sensitivity and false alarm rate and the precision of the seed points in three different stages. Note that the sensitivity and the false alarm rate decrease as we remove potentially erroneous seeds while the precision tends to increase.



(a) Variations in sensitivity & false alarm rate (b) Variations in precision

Figure 3: Seed pruning.

2.3 Region growing

To segment the vessel regions using the selected seeds, we employ a classical region growing method. The region growing process is driven by two hypotheses: (1) the region should grow in the direction of vessels to include areas where no seeds are present and (2) the region should grow in the orthogonal direction until we reach the edges of vessels to provide an accurate segmentation. To achieve these aims, we use the EdgeLoG and the Edgeness operators. Being (x_p, y_p) a seed or a pixel that belongs to the region growing frontier, we analyze its 8-neighborhood to check if to propagate the frontier or not. Being (x_j, y_j) one neighbor of (x_p, y_p) that has not been analyzed previously, we set the pixel as the new frontier if the following constraint is satisfied:

$$\mathrm{EL}(x_j, y_j) < t_{EL} \cup \frac{\mathrm{E}(x_j, y_j) - \mathrm{E}(x_p, y_p)}{||(x_p, y_p) - (x_j, y_j)||} \geq t_{\mathrm{E}} \tag{2}$$

where $t_{EL} \leq 0$ and $t_E \geq 0$ are two algorithm parameters. Equation (2) makes the algorithm grow through the negative valleys of the EdgeLoG signal (see Fig.2(b)) and allows the algorithm to "climb" the Edgeness map, until the ridges are reached (see Fig. 2(c)). A nice property of the proposed method is that the theory behind the EdgeLoG and Edgeness maps suggests a theoretical value for the parameters: $t_{EL} = 0$ and $t_E = 0$. Nonetheless, due to the discrete domain of the image and noise, we set $t_{EL} = -2.5 \cdot 10^{-3}$ and $t_E = 2.5 \cdot 10^{-3}$.

3 VALIDATION

The method has been tested on the DRIVE database. The database contains 40 retinal images divided into a training set and a test set, containing 20 images each. Since our algorithm does not require a training step, we limit our evaluation to the test set. In all the results shown in this paper, the set of scales is $\Sigma = \{0.8, 1.75, 3.5\}$. For the given scales, a value of $k = 1.256$ has been obtained. Table 1 shows the results of our method compared to state-of-the-art algorithms, as reported in the DRIVE website[2]. Our method performs well compared to the state-of-the-art while being less computational expensive and much less complicated than other algorithms.

[2]http://www.isi.uu.nl/Research/Databases/DRIVE/results.php

Table 1: Performance of vessel segmentation methods on the DRIVE database.

Method	Accuracy	Kappa
Human observer	0.9473 (0.0048)	0.7589
Staal [12]	0.9442 (0.0065)	0.7345
Niemeijer [9]	0.9416 (0.0065)	0.7145
Our method	*0.9345 (0.0060)*	*0.6991*
Zana [13]	0.9377 (0.0077)	0.6971
Al-Diri [1]	0.9258 (0.0126)	0.6716
Jiang [4]	0.9212 (0.0076)	0.6399
Martínez-Pérez [8]	0.9181 (0.0240)	0.6389
Chaudhuri [2]	0.8773 (0.0232)	0.3357
All background	0.8727 (0.0123)	0

4 DISCUSSION

The proposed algorithm aims at performing a reasonably good segmentation while limiting the computational cost to the very minimum. The use of linear filtering at different scales allows a quick detection of a certain number of candidate seeds, without the need of computing the Hessian and its eigenvalues for all the pixels in the image. While this could be a minor gain in the algorithm speed for 2D images, it is a huge advantage for 3D data, where the computation of the Hessian and its eigenvalues is extremely costly. To properly deal with possible false positives in the first part of the detection, we employ a Hessian analysis *only on previously selected seeds*; in the proposed experiments the selected seeds are 0.16% of the total image. Moreover, the dominant scale can be easily computed from the EdgeLoG operator, then reducing the computation of eigenvalues to only one scale per seed pixel. The region growing is very efficient and uses a very limited local information, namely the EdgeLoG map and the local gradient of the Edgeness map. The computational cost is mainly dominated by the Gaussian filtering, so that the computational cost is $\mathcal{O}(SN\log N)$, where S is the number of scales and N is the number of image pixels. It is worth mentioning that this part of the algorithm can be easily parallelized.

The method has several interesting properties: (1) the use of the LoG allows having negative EdgeLoG values in presence of bifurcations, so that the region growing can actually fill bifurcation areas; (2) differently than methods that use the eigenvalues of the Hessian matrix, the method does not produce less accurate results in vessels presenting high curvature; (3) finally, thanks to the Edgeness signal, strong edges are suppressed while other methods tend to identify straight edges in 2D images as vessels.

Nonetheless, the method has some limitations to be approached in future research: (1) the use of the Edgeness signal, especially at coarser scales, could potentially mask important information at finer scales, causing e.g. poor segmentation of orthogonal vessels branches whose contrast is much lower than the main vessel; (2) the implemented region growing has no sub-pixel accuracy and very little vessels can be lost in the region growing process due to EdgeLoG and Edgeness discrete lattice.

Finally, a discussion should be devoted to the similarity and differences with the method presented in [8]. Our method does not need the computation of eigenvalues of Hessian matrix. The proposal of the EdgeLoG signal, with appropriate k parameter is totally novel. Despite both methods use a region growing approach, we have theoretical basis on parameter selection, and our method requires much less computational cost: we do not perform region growing of the "background" class.

5 CONCLUSION AND FUTURE WORKS

In this paper we presented an efficient method for segmentation of vessels in medical images. Despite its low computational cost, it demonstrated to perform as good as state-of-the-art algorithms on the DRIVE dataset. Future works encompass: (1) the use of a classifier to learn the constraints for the region growing propagation as a function of local image features; (2) implementation of a 3D version of the method and subsequent testing on CT data.

References

[1] B. Al-Dir, A. Hunter, and D. Steel. An active contour model for segmenting and measuring retinal vessels. *IEEE TMI*, 28(9):1488–97, 2009.

[2] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE TMI*, 8(3):263–269, 1989.

[3] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *Proc. of MICCAI*, volume 1496, pages 130–137, 1998.

[4] X. Jiang and D. Mojon. Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images. *IEEE TPAMI*, 25(1): 131–137, 2003.

[5] C. Kirbas and F. Quek. A review of vessel extraction techniques and algorithms. *ACM CSUR*, 36(2):81–121, 2004.

[6] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trousset. Model-based detection of tubular structures in 3d images. *CVIU*, 80(2):130 – 171, 2000.

[7] T. Lindeberg. Principle for automatic scale selection. Technical report, Royal Institute of Technology, 1998.

[8] M. Martínez-Pérez, A. Hughes, A. Stanton, S. Thom, A. Bharath, and K. Parker. Retinal blood vessel segmentation by means of scale-space analysis and region growing. In *MICCAI*, volume 1679, pages 90–97, 1999.

[9] M. Niemeijer, J. J. Staal, B. van Ginneken, M. Loog, and M. D. Abramoff. Comparative study of retinal vessel segmentation methods on a new publicly available database. In *SPIE MI*, volume 5370, pages 648–656, 2004.

[10] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE TSMC*, 9(1):62–66, 1979.

[11] C. Smets, G. Verbeeck, P. Suetens, and A. Oosterlinck. A knowledge-based system for the delineation of blood vessels on subtraction angiograms. *PRL*, 8(2):113 – 121, 1988.

[12] J. J. Staal, M .D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken. Ridge based vessel segmentation in color images of the retina. *IEEE TMI*, 23:501–509, 2004.

[13] F. Zana and J. Klein. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE TIP*, 10(7):1010–19, 2001.

Diffusion MRI

The application of full body diffusion weighted imaging in oncology

Niall Colgan
n.c.colgan@swansea.ac.uk

School of Medicine,
Swansea University, UK

Abstract

The term PET like MRI refers to the full body diffusion weighted imaging to diagnose metastises in the body. The gold standard of clinical in vivo imaging of metastesis is the utilisation of [^{18}F]FDG radiotracer for PET imaging. FDG is trapped within the metabolically active tumour cells and provides the basis for functional imaging and Maximum Intensity Projections (MIPS) display. MRI full body diffusion weighted imaging has shown promise as a method for displaying similar data to that of PET imaging. The main advantages of this method is that it does not require a radiotracer and that it has a faster clinical through put. However there are barriers to the application of this method in a clinical setting.

1 Introduction

Diffusion weighted imaging is the measurement of the apparent intrinsic self-diffusion property (D) of a fluid according to Brownian motion. Hence D reflects the mobility of the molecules in their microenvironment. Proton NMR (and therefore MRI) can be made sensitive to dynamic displacements of water molecules between 10^{-8} and 10^{-4} m in a timescale of a few milliseconds to a few seconds and these displacements are of the same order of magnitude as cellular dimensions within biological tissues. The water diffusion is affected by the microdynamics of cellular transport between different sub-compartments of the heterogeneous tissue structure as well as by the presence of non-permeable membranes; therefore measurements of D at different parameters, such as the diffusion time, can highlight different properties of the tissues morphology. By entangling the rate of diffusion with relaxation time provides a unique image contrast of the diffusion coefficient referred to as the apparent diffusion coefficient (ADC). The main application of DWI in clinical and research setting are the evaluation of acquired brain injury including stroke and traumatic brain injury[1-6].

However ADC of cancer cells and tumour sites is reduced due to increased cell diameters and the dense cellular composition in comparison to surrounding tissue. Therefore, the imaging sequence should be able to produce similar outputs to PET imaging in the study of secondary metastases which are faster and without the need for contrast agents.

2 Current application and limitations

The application of fully body scanning with Diffusion weighted imaging has only recently evolved into a usable structure. The main limitation has been the development of a sequence that can be utilised in a free breathing as breath hold scanning has SNR limitations due to short acquisition times eliminating the possibility of thin slice acquisitions with good SNR for PET like MIPS display. The first protocol allowing free breathing DWIBS was developed on a Philips™ 1.5 T scanner [7]. The protocol relied on a STIR with EPI in free breath with the signal average is performed on the reconstructed image and not in K-space. This aims to average out the motion artefact.

The other vendors have produced similar imaging sequences like Siemens™ REVEAL however GE™ have yet to produce a DWI full body sequence. The main limitations of the sequence design and the future bases for future research:

Scientific limitations

1. Lack of understanding of DW-MRI at a microscopic level
2. False positives, any densely packed tissue mass (e.g. cyst) will look like a tumour
3. No accepted standards for measurements and analysis
4. Multiple data acquisition protocols depending on body part and usage of data
5. Qualitative to quantitative assessments
6. Multi-exponential decay components which affect the calculated ADC values

Vendor specific

1. Rapid evolution of body imaging protocols and coils
2. Incomplete validation and documentation of reproducibility
3. Divergent nomenclature and symbols
4. Lack of working methodologies, accepted quality assurance
5. No (QA) standards, and physiologically realistic phantoms
6. Requires improving measurement and analysis methods with repeatability, and reproducibility.

3 Conclusion

The main barrier to the application of this method in a clinical setting is the possibility of false positives of non-cancerous lesions or cysts in patients. This is the main challenge to the application of this method.

1. Colgan, N.C., *Quantitative assessment of focal brain injury from impact*. 2009, University College Dublin, Ireland: Ph.D. thesis.
2. Nagesh, V., et al., *Time course of ADCw changes in ischemic stroke: beyond the human eye!* Stroke, 1998. **29**(9): p. 1778-1782.
3. Romero, J.M., et al., *Diffusion MR imaging of acute ischemic stroke.* Neuroimaging Clinics of North America, 2002. **12**(1): p. 35-53.
4. Warach, S., J.F. Dashe, and R.R. Edelman, *Clinical outcome in ischemic stroke predicted by early diffusion-weighted and perfusion magnetic resonance imaging:*

a preliminary analysis. Journal of Cerebral Blood Flow & Metabolism, 1996. **16**(1): p. 53-59.

5. Huisman, T., et al., *Diffusion-weighted imaging for the evaluation of diffuse axonal injury in closed head injury.* Journal of Computer Assisted Tomography, 2003. **27**(1): p. 5.
6. Bammer, R., *Basic principles of diffusion-weighted imaging.* European journal of radiology, 2003. **45**(3): p. 169-184.
7. Takahara, T., et al., *Diffusion weighted whole body imaging with background body signal suppression (DWIBS): technical improvement using free breathing, STIR and high resolution 3D display.* Matrix, 2004. **160**(160): p. 160.

White Matter Models of *In Vivo* Diffusion MRI Human Brain Data: A Statistical Ranking

Uran Ferizi[12]
u.ferizi@cs.ucl.ac.uk
Eleftheria Panagiotaki[1]
e.panagiotaki@cs.ucl.ac.uk
Torben Schneider[2]
t.schneider@ucl.ac.uk
Claudia Wheeler-Kingshott[2]
c.wheeler-kingshott@ucl.ac.uk
Daniel Alexander[1]
d.alexander@ucl.ac.uk

[1] Centre for Medical Image Computing
Department of Computer Science
University College London
London, UK

[2] NMR Research Unit
Department of Neuroinflammation
Institute of Neurology
University College London
London, UK

Abstract

Diffusion MRI (Magnetic Resonance Imaging) Microstructure Imaging provides a unique non-invasive probe into the microstructure of biological tissue. However, it relies on a mathematical model relating microscopic tissue features to the MR signal. This work aims to determine which models of diffusion MRI are best at describing the signal from *in vivo* brain white matter. The assumption of Gaussian diffusion in the most commonly used model, the Diffusion Tensor, oversimplifies the diffusive behaviour of water in complex media and is known to break down for relatively large diffusion weights (b-values). Recent work shows that three-compartment models, incorporating restricted intra-axonal compartments, glial compartments and hindered extra-cellular diffusion, are best at explaining multi b-value data sets from fixed brain tissue. Here we perform a similar experiment using *in vivo* human data to avoid, and evaluate, the effects of the fixation process. We compare one-, two- and three-compartment models, and rank them using two standard model selection criteria. Results show that, as with fixed tissue, three-compartment models explain the data best, although simpler models emerge from the *in vivo* data. Both changes in diffusion behaviour from fixation and the lower gradient strengths available *in vivo* are likely to contribute to the difference. The full ranking assists the choice of model and imaging protocol for future brain microstructure imaging.

1 Introduction

Diffusion MRI measures the water dispersion in biological tissue and can therefore be used to probe the microstructure. Though useful in other tissue types, this technique is most often applied in the brain, especially where parallel fibres restrict water mobility anisotropically, thus providing putative measures of white matter integrity and connectivity.

Currently, the standard model for imaging diffusion in tissue is the Diffusion Tensor (DT) [7], which assumes a trivariate Gaussian dispersion pattern. Derived indices, such as Mean

Diffusivity (MD) and Fractional Anisotropy (FA), can correlate with major tissue damage, but lack the sensitivity and specificity to provide indices such as axon radius, density, orientation, dispersion and permeability, which potentially give much greater insight into tissue architecture and pathology. These can be estimated by applying more complex models to richer data sets [2, 3, 4, 5, 6, 12, 14, 15].

Stanisz et al. [14] pioneered a multi-compartment representation of the separate diffusive processes in nerve tissue. The Ball-Stick model [8] is the simplest possible two-compartment model with restricted axonal diffusion and isotropic extra-axonal diffusion. The CHARMED model (Composite Hindered and Restricted Model of Diffusion) [4] is a similar two-compartment model that allows anisotropic Gaussian Diffusion in the 'hindered' extra-cellular space and non-zero-radius cylindrical intra-cellular diffusion. The AxCaliber technique [5] extended the CHARMED model to estimate the distribution of axon diameters. Barazany et al. [6] demonstrated the approach *in vivo* on a rat's Corpus Callosum (CC), adding a third free-diffusion compartment to account for Cerebrospinal Fluid (CSF) contamination. Alexander et al. [2] used a similar model to derive orientationally invariant indices of axon diameter and density.

Increasing model complexity can potentially lead to overfitting and false parameter estimation. Recently, Panagiotaki et al. [12] compared 47 diffusion MRI models using data from fixed White Matter (WM) of rats, using the Bayesian Information Criterion (*BIC*). They proposed a taxonomy of one-, two- and three- compartment models as a combination of: i) a hindered/extra-cellular diffusion via a Tensor (an ellipsoid), a Zeppelin (a cylindrically symmetric ellipsoid) or a Ball (an isotropic sphere); ii) a restricted/intra-cellular diffusion via a Stick (an oriented line with zero radius), a Cylinder (a Stick with non-zero constant radius) or Gamma Distributed Cylinders (not included in this work); iii) an isotropically restricted component, representing glial cells or isotropically oriented axons, modelled by a Dot ("stationary" water), a Sphere (spherically restricted), Astrosticks (zero radius lines isotropically spread in space) or Astrocylinders (similarly to the Astrosticks, but with non-zero radii). They concluded that the three-compartment models perform best and DT worst.

Here we repeat the experiment in [12] using *in vivo* data from the human brain's CC and the software used in [12] as part of the Camino toolkit [10]. We acquire a rich data set with many combinations of diffusion and gradient times and gradient magnitudes to fit the models to as wide a fraction of the measurement space as possible. We test the quality-of-fit of models to the data using *BIC*, and confirm with the Akaike Information Criterion (*AIC*).

2 Methods

This section describes the acquisition protocol for our rich data set and outlines the preprocessing we perform to obtain a set of measurements for model fitting. It then details the fitting procedure and the techniques we use for model selection and ranking.

2.1 DW/DT Acquisition

Using a 3T Phillips scanner, we scan a 30-year old healthy man, using three gradient-encoding directions: one approximately along CC main fiber direction and two in its perpendicular plane. The images consist of 8x4mm sagittal slices, an image size of 64x64 and in-plane resolution of 2mm x 2mm.

Using a Pulsed-Gradient-Spin-Echo sequence, we probe gradient strengths |G| = 30, 40, 50, 60 mT/m, gradient times δ = 5, 15, 25ms, and diffusion times Δ ranging in steps of 10ms, from 20 to 100ms. This produces a total of 63 diffusion weightings, excluding any combinations where δ>Δ, with maximum weight b=8,300s/mm^2. As the echo times vary across different acquisition parameters, in addition to every diffusion-weighted acquisition we obtain a corresponding non-diffusion-weighted (b=0) one. We also perform a separate HARDI (High Angular Resolution Diffusion Imaging) acquisition with the same image resolution, 32-gradient encoding directions and b=711s/mm^2. The total acquisition time is 2.5hrs.

2.2 Data Preprocessing

As in [12], we first fit the DT model to the HARDI data and find the principal direction to identify voxels containing fibres well aligned with the assumed fibre direction. We take all voxels with FA above a threshold and principal eigenvector within a small angle of tolerance of the assumed fiber direction and average them to give a single data set to which the model can be fitted. Here, we set the FA threshold at 0.5 and repeat the experiment with three different angular thresholds, 2°, 5° and 10°, to establish the effect of orientation dispersion. For b>3,500s/mm^2, the parallel direction signal reaches the noise floor, so we exclude these data, thus leaving us with 171 DW (and 63 unweighted) images. The SNR is about 20. The signal at each DW is normalised by the corresponding diffusion-unweighted measurement with the same echo time, to remove T2 effects, before the fitting.

2.3 Model Fitting

We fit 32 models of diffusion to the signal, using the open software tool Camino [10]. We fit in two stages: after an initial run of 1,000 random starting points, we extract the parameters that produce the minimum objective function. We then execute another 1,000 runs from starting points at small random perturbations from the first minimum. This ensures that we avoid local minima and obtain the best fit parameters. The offset-Gaussian noise model is used, as in [12], to construct the objective function to minimize (*RES*). This objective function accounts for bias introduced by the Rician noise inherent in the data in a simplistic way [11] that is more numerically stable than a full Rician log-likelihood objective function.

2.4 Model Selection

Model selection criteria, as introduced in [13]

$$BIC = -2\log(L) + K\log(N) \tag{1}$$

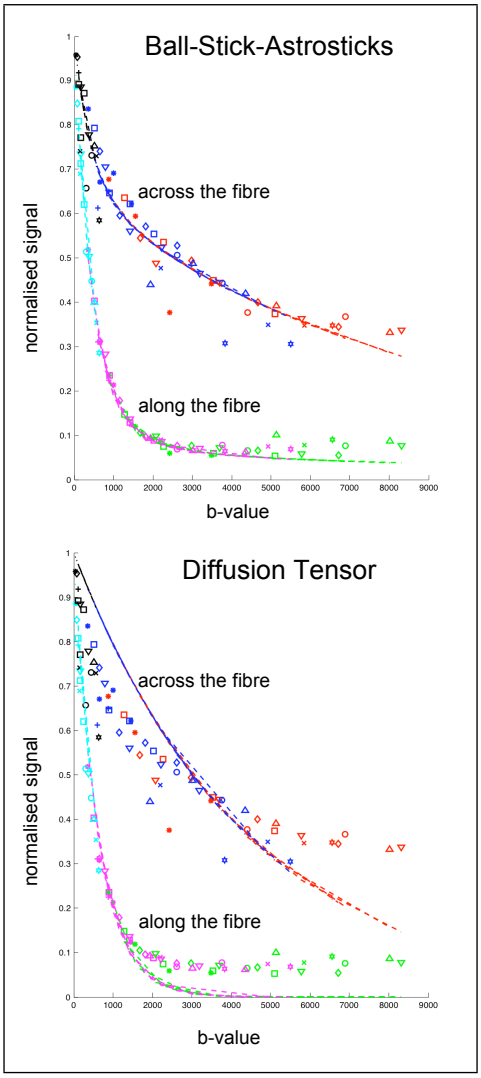
where L is the likelihood of obtaining the present data given the model, K is the number of free parameters and N is the number of measurements, or in [1]

$$AIC = -2\log(L) + 2K + \frac{2K(K+1)}{(N-K-1)} \tag{2}$$

quantify the trade-off between complexity and goodness-of-fit to identify the simplest model that explains some data. The lower the scores, the more predictive the model is. The last term in the *AIC* is the adaptation suggested in [9] of the original *AIC* for cases where the number of measurements is not large (deemed to be smaller than about 40 times the number of free parameters). In general, the *AIC* is less conservative in penalising complexity.

Parameters	Model	2° deviation BIC order	2° deviation AIC order	2° deviation RES	5° deviation BIC order	5° deviation AIC order	5° deviation RES	10° deviation BIC order	10° deviation AIC order	10° deviation RES
6	Ball.Stick.Ast	**1**	1	149.05	**1**	1	94.54	**1**	1	74.41
6	Ball.Stick.Dot	**2**	6	151.64	**2**	6	96.44	**2**	6	76.61
7	Ball.Stick.Acl	**3**	2	149.05	**5**	2	94.29	**6**	2	73.73
7	Ball.Cylin.Acl	**4**	3	149.06	**3**	3	94.23	**4**	3	73.52
7	Ball.Cylin.Ast	**5**	4	149.06	**4**	4	94.28	**5**	4	73.61
7	Zepp.Stick.Ast	**6**	5	149.07	**6**	5	94.48	**3**	5	73.48
5	Ball.Stick	**7**	21	162.38	**7**	21	106.82	**9**	21	87.00
7	Zepp.Stick.Dot	**8**	10	151.64	**8**	10	95.97	**7**	10	74.83
7	Ball.Cylin.Dot	**9**	11	151.65	**9**	11	96.16	**8**	11	75.79
7	Ball.Stick.Sph	**10**	12	151.86	**10**	12	96.41	**10**	12	76.51
8	Zepp.Stick.Acl	**11**	7	149.08	**13**	7	94.24	**13**	7	72.99
8	Zepp.Cylin.Acl	**12**	8	149.09	**11**	8	94.19	**12**	8	72.89
8	Zepp.Cylin.Ast	**13**	9	149.14	**12**	9	94.23	**11**	9	72.89
6	Ball.Cylin	**14**	24	162.38	**15**	24	106.60	**17**	24	86.33
6	Zepp.Stick	**15**	25	162.39	**14**	25	105.80	**15**	25	85.44
8	Zepp.Cylin.Dot	**16**	14	151.64	**16**	14	95.72	**14**	14	74.21
8	Ball.Cylin.Sph	**17**	15	151.76	**18**	15	96.16	**18**	15	75.79
8	Zepp.Stick.Sph	**18**	16	151.77	**17**	16	95.92	**16**	16	74.62
9	Tens.Stick.Ast	**19**	13	149.09	**19**	13	94.46	**19**	13	73.43
9	Tens.Stick.Dot	**20**	17	150.96	**23**	17	96.27	**23**	17	76.30
7	Zepp.Cylin	**21**	27	162.39	**20**	27	105.57	**20**	27	84.85
7	Bizepp.	**22**	28	162.39	**21**	28	105.80	**22**	28	85.44
9	Zepp.Cylin.Sph	**23**	20	153.48	**22**	20	95.72	**21**	20	74.20
10	Tens.Cylin.Acl	**24**	18	149.09	**26**	18	94.21	**25**	18	72.97
10	Tens.Stick.Acl	**25**	19	149.14	**25**	19	94.16	**24**	19	72.79
8	Tens.Stick	**26**	30	162.41	**24**	30	105.21	**27**	30	84.68
10	Tens.Cylin.Dot	**27**	22	151.72	**28**	22	95.65	**28**	22	74.17
10	Tens.Stick.Sph	**28**	23	152.30	**29**	23	95.86	**29**	23	74.58
9	Tens.Cylin	**29**	31	162.56	**30**	31	105.09	**30**	31	84.48
10	Tens.Cylin.Ast	**30**	29	157.64	**27**	29	95.40	**26**	29	73.18
11	Tens.Cylin.Sph	**31**	26	152.29	**31**	26	95.69	**31**	26	74.17
7	Tensor	**32**	32	616.71	**32**	32	477.12	**32**	32	449.85

(a)



(b)

Figure 1: a) Models ranked for the 2°/5°/10° fibre deviation allowances. Adjacent are the raw scores for their respective Objective Function Residue (*RES*). The Ball-Stick-Astrosticks does best across both *BIC* and *AIC*, the DT worst. (Abbrev.: Tens.=Tensor; Zepp.=Zeppelin; Cylin.=Cylinder; Ast.=Astrosticks; Acl.=Astrocylinders; Sph.=Sphere). b) Matching the raw signal and the signal synthesized from the model parameters for the 2° threshold data set. The Ball-Stick-Astrosticks model fits the data well; the Tensor (DT) fits poorly.

3 Results

The ranking of the models is provided in the table of Fig.1, separately for 2°, 5° and 10° angular threshold data sets, for both *BIC* and *AIC* . The models are listed according to the *BIC* performance in the 2° data set. We also give the raw scores for the *BIC*, *AIC* and *RES*.

The three compartment models describe the data best. The single compartment model, the DT, performs worst across all criteria. The most complex model, the Tensor-Cylinder-Astrocylinders, gives small *RES*, but its complexity is penalised under *AIC*, more under *BIC*.

In the extra-cellular space, the Ball, the Zeppelin and the Tensor perform equally well, judged on their *RES* score, but the simpler Ball scores better on *BIC*. The intracellular Stick does best with *BIC*, but Cylinders perform well too, especially across all 3 scores (e.g. Ball-Cylinder-Astrocylinders). Across the third compartments, the Sphere performs worst, partly because of its model complexity.

The *RES* scores decrease markedly as the orientation threshold increases from 2° to 5°, but at a smaller rate from 5° to 10°. This suggests that the models fit increasingly better as the number of voxels averaged increases.

The plots of Fig.1 show the fit of two models, the best and the worst from the ranking, to the data, revealing the limitations of the overly simple single-compartment DT model. The DT model cannot capture the shape of the perpendicular signal decay because it does not model restriction. The three-compartment model, Ball-Stick-Astrosticks, fits both the parallel and perpendicular direction signal decay closely.

4 Discussion

By using a wide range of diffusion weights and times, we showed that standard models from the literature, such as the DT and Ball-Stick, are too simple to describe diffusion in the brain. Three compartments are required to explain the diffusion in white matter *in vivo*. These models fit the data best, with very similar precision, and the simpler members of this class are favoured under the BIC ranking. Further work with high angular resolution data sets will shed more light on this.

The *BIC* for the Cylinder models are only marginally lower than those of Stick, supporting a future use of cylinder models for estimating axon diameter *in vivo*, as in [3, 6, 15].

For the third compartment, the Astrosticks or Astrocylinders show an advantage over the Sphere and Dot, which were favoured for fixed tissue in [12]. This could be because of intrinsic differences between live and fixed tissue, or differences in the imaging protocol, since we use much lower gradient strength. Intrinsic tissue differences are a likely cause; Alexander in [3] suggests fixation may make membranes 'stickier', giving rise to slow moving water well captured by the Dot model.

Minor variations in the dispersion of fibre direction in the data set affect fitting quality, since the models we test here capture either no dispersion or a fully uniform one. We intend to consider localised dispersion in future work using, for example, the models in [15]. The Astrosticks and Astrocylinders models may improve the fit simply by accommodating small amounts of fibre orientation dispersion even at the 2° threshold, or may genuinely capture a small population of fibres with uniformly distributed orientation.

We also emphasize that the choice of models our analysis suggests is not appropriate for existing sparse data sets such as off-the-shelf single shell HARDI data which only support simple models. Rather, these results inform the choices for future *in vivo* microstructure imaging once we identify the right model. Experiment design techniques such as [2] can determine economical protocols. We emphasize that the protocol we use here is designed specifically for model selection rather than large scale application.

As, intentionally, we conservatively selected the part of the CC that is mostly homogeneous, with little fibre crossing or CSF contamination, other models may perform better across or away from the CC (e.g. a CSF pool as in [6]). Future work will include other models such as a distribution of pore sizes [5, 12] and orientation [15].

5 Acknowledgements

EPSRC kindly funds the Research Studentship of U.Ferizi. We are grateful to the Multiple Sclerosis Society of Great Britain and Northern Ireland for supporting the NMR unit.

This work was undertaken at UCLH/UCL who received a proportion of funding from the Department of Health NIHR Biomedical Research Centres funding scheme.

References

[1] H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.

[2] D.C. Alexander. A general framework for experiment design in diff. MRI and its application in measuring direct tissue-microstructure features. *MRM*, 60(2):439–448, 2008.

[3] D.C. Alexander, P.L. Hubbard, M.G. Hall, E.A. Moore, M. Ptito, G.J.M. Parker, and T.B. Dyrby. Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage*, 52(4):1374–1389, 2010.

[4] Y. Assaf, R.Z. Freidlin, G.K. Rohde, and P.J. Basser. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. *Magnetic Resonance in Medicine*, 52(5):965–978, 2004.

[5] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, and P.J. Basser. Axcaliber: a method for measuring axon diameter distribution from diff. MRI. *MRM*, 59(6):1347–1354, 2008.

[6] D. Barazany, P.J. Basser, and Y. Assaf. In vivo measurement of axon diameter distribution in the corpus callosum of rat brain. *Brain*, 132(5):1210, 2009.

[7] P.J. Basser, J. Mattiello, and D. LeBihan. MR diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267, 1994.

[8] TEJ Behrens, MW Woolrich, M. Jenkinson, H. Johansen-Berg, RG Nunes, S. Clare, PM Matthews, JM Brady, and SM Smith. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *MRM*, 50(5):1077–1088, 2003.

[9] K.P. Burnham and D.R. Anderson. Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research*, 28(2):111–119, 2001.

[10] PA Cook, Y. Bai, S. Nedjati-Gilani, KK Seunarine, MG Hall, GJ Parker, and D.C. Alexander. Camino: Open-source diffusion-MRI reconstruction and processing. In *14th scientific meeting of the ISMRM*, volume 2759, 2006.

[11] D.K. Jones and P.J. Basser. "Squashing peanuts and smashing pumpkins": How noise distorts DW-MR data. *MRM*, 52(5):979–993, 2004.

[12] E. Panagiotaki, T. Schneider, B. Siow, M.G. Hall, M.F. Lythgoe, and D.C. Alexander. Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. *NeuroImage*, 2011.

[13] G. Schwarz. Estimating the dimension of a model. *The A. of Stat.*, 6(2):461–464, 1978.

[14] G.J. Stanisz, G.A. Wright, R.M. Henkelman, and A. Szafer. An analytical model of restricted diffusion in bovine optic nerve. *MRM*, 37(1):103–111, 1997.

[15] H. Zhang, P.L. Hubbard, G.J.M. Parker, and D.C. Alexander. Axon diameter mapping in the presence of orientation dispersion with diffusion MRI. *NeuroImage*, 2011.

Laplace–Beltrami Regularization for Diffusion Weighted Imaging

Bartosz P. Neuman[1]
bartosz.neuman@nottingham.ac.uk
Christopher Tench[2]
christopher.tench@nottingham.ac.uk
Li Bai[1]
bai.li@nottingham.ac.uk

[1] School of Computer Science
University of Nottingham
Nottingham, UK

[2] Academic Division of Clinical Neurology
University of Nottingham
Nottingham, UK

Abstract

A popular method of smoothing the diffusion signal is based on a Tikhonov regularization with Laplace–Beltrami operator. In this study a systematic Monte-Carlo study of a regularization parameter selection is presented. By minimizing the difference between synthetic and reconstructed shapes the best parameter values are found. Those parameters are then tested on a real MRI image. The fibre orientation distribution functions obtained with new regularization parameters are more stable and reliable.

1 Introduction

Magnetic Resonance Imaging (MRI) is a medical imaging technique which allows for *in vivo* visualisation of internal structures in detail. It is especially sensitive to different soft tissues, producing a good contrast between them. Over the course of past decades various types of MRI applications have been introduced, some of which are linked to the measurement of water diffusion. Diffusion weighted imaging of the brain uses the displacement of water molecules in the tissue to introduce contrast to images [8]. The diffusion of water can be used to infer white matter fibre orientations and to study brain connectivity in vivo [10].

One of the major areas of research in diffusion imaging is the reconstruction of the distribution of water diffusion. Using the diffusion weighted imaging a single scalar diffusivity, which depends on the direction of the applied magnetic field gradient (called the diffusion sensitizing gradient), is calculated. The signal is usually approximated with a 3D function, which depending on the processing of the signal, results in a diffusion profile, diffusion orientation distribution function (ODF), or fibre orientation distribution function (FODF). Unfortunately, as the signal is heavily corrupted by noise the reconstructed signal shapes are spiky and have to be smoothed.

An elegant and straightforward way of obtaining a more regular function is to use a Tikhonov regularization during the function fitting process. One of the most successful and widely used regularizations is based on the Laplace–Beltrami operator [4]. But despite its long use, there have been almost no study regarding the selection of the regularization parameter, which relates to the strength of smoothing applied.

In the original paper [4], an L-curve analysis [7] was used to establish the optimal parameters for a very specific case ($b = 3000$ s/mm^2 and SNR $= 35$). Depending on the number of fibres within a voxel it was reported that the optimal parameter is 0.308 for one, 0.006 for two, and 0.0155 for 3 fibre crossing. 0.006 was considered a good compromise between a regularization and angular resolution, and recommended for general use. Another method, based on a generalized cross validation (GCV) used an iterative approach to automatically select the regularization parameter [9]. Later on, it was shown that both L-curve and GCV methods provide similar parameters [3], with L-curve being more stable.

The choice of the regularization parameter based on either L-curve or GCV is a compromise between preserving the data and improving the condition of the design matrix. This, due to the data being corrupted by noise, can result in an under-smoothed solution (noise spikes will be treated as a valid data). Intuitively, instead of deciding on the regularization parameter based on the goodness of fit to the noisy, measured data, one should base the decision on the unfitness to the ground truth. Additionally, depending on the objective function (best fit of diffusion profile, ODF, FODF) the regularization parameter should differ.

In this paper a thorough investigation is conducted to select regularization parameters, using such a ground truth validation. Tests are performed for a range of diffusion weighting factors, SNRs and objective functions. The findings are finally presented on a real image.

2 Methods

2.1 Synthetic Data Generation

The synthetic diffusion signal was created using the multiple tensor model [6]:

$$S_r = S_0 \sum_{k=1}^{N} v_k \exp(-b\vec{g}_r^{\mathrm{T}} \mathbf{D}_k \vec{g}_r) \,, \tag{1}$$

where S_r is the amplitude of the measured signal with the diffusion sensitizing gradient applied at direction r, S_0 with no diffusion sensitizing gradients applied, b a diffusion-weighting factor, $\vec{g}_r$ the direction of the diffusion sensitizing gradient, the $\mathbf{D}_k$ a diffusion tensor of the kth fibre, and v_k a volume fraction ($\sum_{k=1}^{N} v_k = 1$). Both the number of the fibre crossings and the volume fractions were selected randomly. Additionally, on average half of the simulated voxels had a partial isotropic environment present. The signal was corrupted by adding a complex Gaussian noise and taking the magnitude.

2.2 Spherical Harmonics Analysis

Spherical harmonics (SH) allow for a non-parametric analysis of the diffusion signal, and can be seen as the extension of Fourier basis functions to the sphere. The basis functions, defined as a solution to Laplace equation, are given by:

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} P_l^m(\cos\theta)\exp(im\phi) \,, \tag{2}$$

with P_l^m being the associated Legendre function of band (order) l and degree m, θ a colatitude, and ϕ longitude. For practical reasons a single integer $j(l,m) = (l^2+l+2)/2+m$ to enumerate the spherical harmonic series (SHS) was often used.

The diffusion signal, measured using high angular resolution diffusion imaging protocol [11], can be approximated with a truncated series of SH basis functions:

$$S(\theta_i,\phi_i) = \sum_j c_j Y_j(\theta_i,\phi_i)\,, \tag{3}$$

where c_j are SHS coefficients. Due to antipodal symmetry of diffusion, the odd order basis functions do not appear in Equation 3.

The parameters $\mathbf{c}$ can be found using classical linear regression by minimizing the Euclidean norm:

$$\arg\min_{\mathbf{c}} \|\mathbf{Yc}-\mathbf{S}\|^2\,, \qquad \mathbf{c} = (\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T\mathbf{S}\,, \tag{4}$$

where $\mathbf{Y}$ is a spherical harmonics design matrix, and $\mathbf{S}$ a vector with diffusion weighted measurements. To give preference to a solution with certain desirable properties (like smoothness), the regularization term is included:

$$\arg\min_{\mathbf{c}_\lambda} \|\mathbf{Yc}_\lambda-\mathbf{S}\|^2+\lambda^2\|\Gamma\mathbf{c}_\lambda\|^2\,, \qquad \mathbf{c}_\lambda = (\mathbf{Y}^T\mathbf{Y}+\lambda^2\Gamma^T\Gamma)^{-1}\mathbf{Y}^T\mathbf{S}\,, \tag{5}$$

with Γ a chosen Tikhonov matrix and λ^2 a parameter controlling the strength of regularization. For $\lambda^2 = 0$ this reduces to the classical least squares solution. The Tikhonov regularization using Laplace–Beltrami operator, which is a natural measure of smoothness for functions defined on a sphere, is achieved by using a diagonal matrix $\Gamma^T\Gamma$ with $l_j^2(l_j+1)^2$ along the diagonal (l_j being the order associated with the jth coefficient).

In order to infer the fibre orientation within a voxel the diffusion signal needs to be represented as either diffusion or fibre ODF. Spherical harmonics allow for an efficient analytical computation of both, ODF using a Funk-Hecke theorem [1] and FODF using a spherical deconvolution transform [5].

2.3 Optimal Regularization Parameter

To find the optimal regularization parameter, a case by case evaluation was performed (through exhaustive search). For a set of diffusion weighting factors and SNRs (both feasible and infeasible combinations), $\lambda^2 \in [0, 0.5]$ was tested. No limits on minimal crossing angle or volume fractions was placed. The best parameter was the one minimizing (on average) the difference between synthetic and reconstructed signal:

$$\arg\min_{\lambda} \|\mathbf{WA}(\mathbf{Y}^T\mathbf{Y}+\lambda^2\Gamma^T\Gamma)^{-1}\mathbf{Y}^T\hat{\mathbf{S}} - \mathbf{WA}(\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T\mathbf{S}\|^2\,, \tag{6}$$

where both synthetic and reconstructed signal was approximated using 8th order SHS. It is possible to search for the optimal regularization parameters based on a different fitness measure, like correlation coefficient or angular error of estimated fibres. In those cases, the observed trends and obtained results should be similar to the ones presented here.

$\mathbf{W}$ and $\mathbf{Y}$ are both 8th order SH design matrices, the former is based on a dense spherical sampling (1024 diffusion encoding gradients, to evaluate the fitness), while the latter on a sparse, physically feasible measurement (60, diffusion encoding gradients, to find the SH coefficients), $\hat{\mathbf{S}}$ is a noise corrupted signal, and $\mathbf{S}$ a ground-truth. The diagonal matrix $\mathbf{A}$ transforms the SH expansion of the signal to either ODF or FODF. For $\mathbf{A} = \mathbf{I}$ this will in effect minimize the difference in diffusion profiles itself.

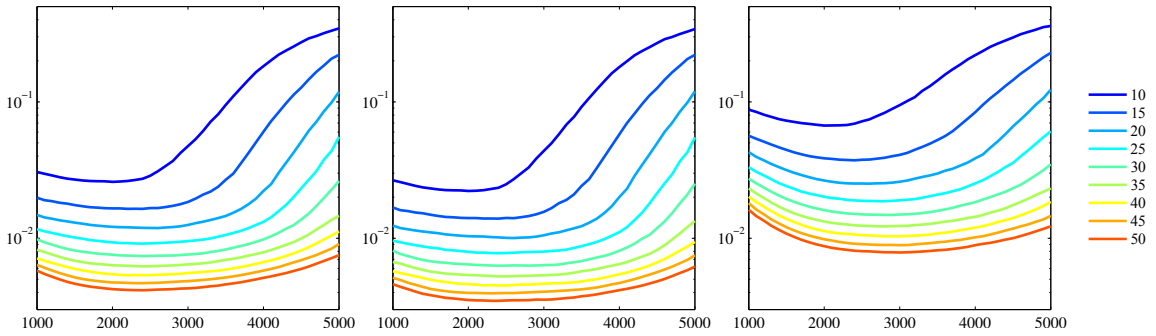


Figure 1: Regularization parameters (log-scale) as a function of diffusion weighting factor. Recommended parameter values that maximize the accuracy of (from top left) diffusion signal reconstruction, ODF, and FODF computation.

The Equation 6 can be split into two parts: reconstructed (left) and synthetic (right) signal. In both cases, the resolution of the signal is raised (using matrix $\mathbf{W}$) and transformed into other representation (ODF/FODF, using matrix $\mathbf{A}$) if necessary. The algorithm for creating the phantom signal and analysing the regularization fitness can be summarized as follows:

1. For a given diffusion weighting factor b and signal to noise ratio $1/\sigma^2$ do:
2. Select a random n (1-3) fibre crossing with random volume fractions $\vec{v}$.
3. Generate a synthetic signal S using a multi-tensor model.
4. Generate a noise corrupted signal $\hat{S} = \sqrt{(S+N(0,\sigma^2))^2+N(0,\sigma^2)^2}$.
5. Investigate the fitness function for various regularization parameters λ^2.
6. Repeat from 2 for 10000 times.

3 Results and Discussion

The optimal parameters found through this research are presented in the Figure 1, where λ^2 values for different SNR cases are plotted as a function of b-value. As expected, the parameters vary depending on both diffusion weighting that is applied and SNR.

With a low diffusion weighting applied it is possible to smooth the profile without the loss of information and accuracy. The signal becomes sharper with the increasing b-values, and the smoothing that can be applied during reconstruction decreases. This trend, however, at certain point reverses, and the recommended smoothing starts to increase. With higher diffusion weighting applied the frequency range of the signal and noise become similar, and the smoothing will remove the real signal and noise corruption indiscriminately.

With a noisier signal it is necessary to apply stronger smoothing, and prevent the resulting function from adapting to noise. At the same time, the regularization parameter should not be too large, as otherwise it will suppress the measured signal. In the simulations, a maximal parameter was set to 0.5.

The recommended parameters for reconstructing the diffusion signal and ODF are very similar (Figure 1, left and middle). At the point of interest ($b = 3000$ s/mm^2 and SNR $= 35$) both are close to 0.006 that was found analytically in [4]. In the case of FODF though, the recommended parameters are twice larger (0.012 at the point of interest).

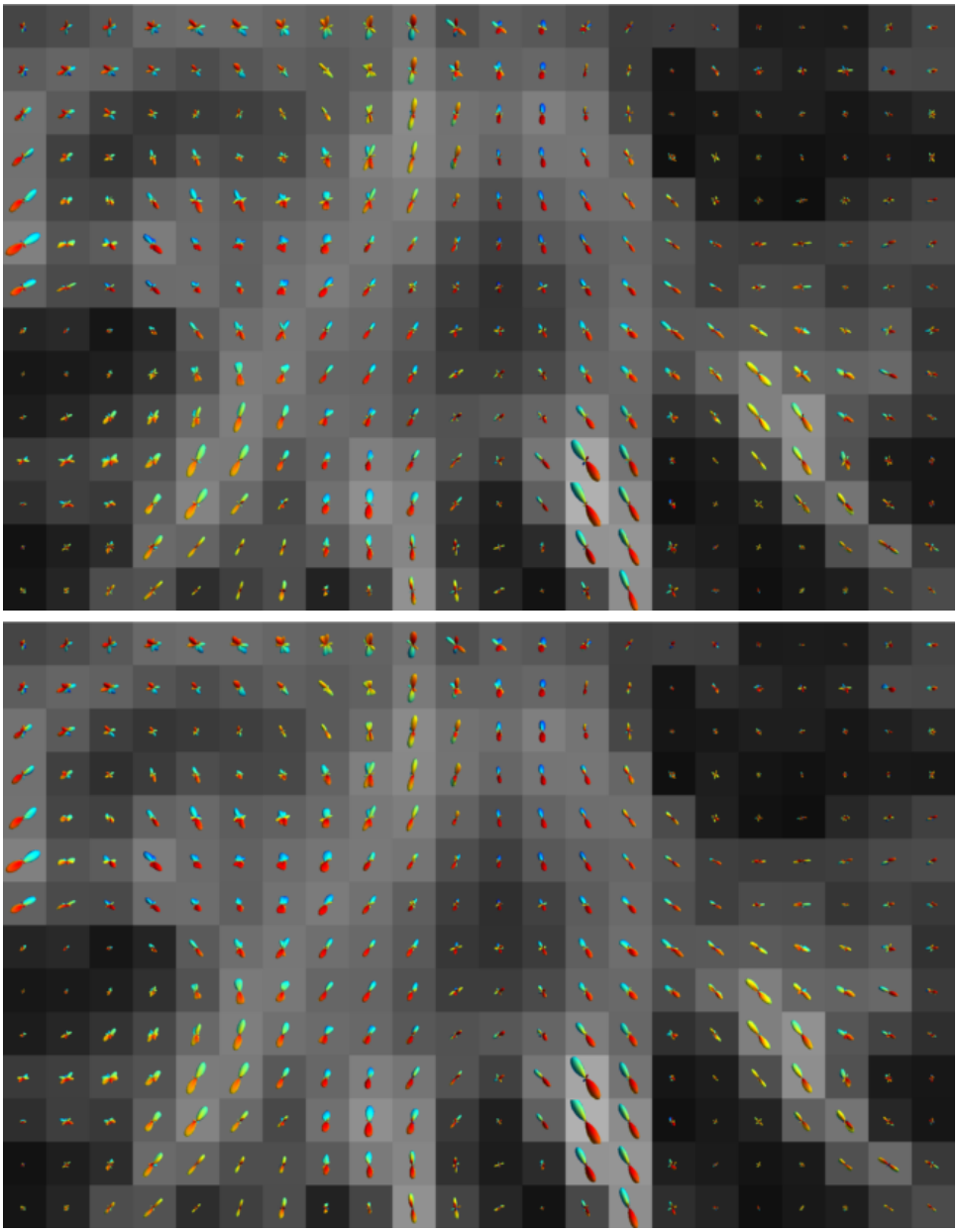


Figure 2: FODFs computed over a region of interest using two regularization parameters: $\lambda^2 = 0.006$ (top) and $\lambda^2 = 0.012$ (bottom).

The best regularization parameter were found for one, two, and three fibre crossings separately. The regularization parameter for accurate one fibre FODF estimation were approximately three times larger, and three fibre crossing were twice of the parameters presented in the Figure 1. This behaviour is similar to the one observed by Descoteaux [4], and likewise, we decided to use the two-fibre results to select the global parameter value.

Finally, a region of interest from a whole-brain scan of a healthy male subject was used to visualize the influence of a stronger regularization. The image was obtained using a single-shot, spin-echo, echo-planar, diffusion-weighted sequence in a Philips 3T Achieva clinical imaging system[1]. Figure 2 shows FODFs computed over the same region regularized with two parameters – 0.006 and 0.012, our optimal FODF reconstruction parameter for $b =$

[1]Acquisition matrix 112x112 with in-plane resolution 2×2 mm^2; 52 slices with a thickness of 2 mm; $b =$ 3000 s/mm^2; $TE = 72$ ms; $TR = 15292$ ms; 61 evenly spaced diffusion weighting directions [2]; six $b = 0$ s/mm^2 images acquired and averaged

3000 s/mm^2 and SNR = 35 (Figure 1, right). The difference between the two is very subtle. The stronger regularization produced less ringing in the computed FODFs and suppressed some of the weaker false peaks. At the same time there was no significant loss of the angular resolution. The increase in FODFs stability is considered beneficial.

4 Conclusion

A systematic investigation of regularization parameter selection has been performed. Optimal parameter values for diffusion profiles, and diffusion and fibre ODFs were obtained by minimizing the difference between synthetic and reconstructed signal shapes. The findings not only corroborate previous research results but also extend them, providing a more complete picture of the parameter selection problem. While the regularization parameters for diffusion signal and ODF reconstruction are confirmed to be similar to those found in the previous research, the FODF reconstruction allows for regularisation parameters several times larger. Using the fine tuned parameter more stable and reliable FODFs were obtained.

References

[1] A.W. Anderson. Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54:1194–1206, 2005.

[2] P.A. Cook, M. Symms, P.A. Boulby, and D.C. Alexander. Optimal acquisition orders of diffusion-weighted MRI measurements. *Journal of Magnetic Resonance*, 25:1051–1058, 2007.

[3] I.J. Day. On the inversion of diffusion NMR data: Tikhonov regularization and optimal choice of the regularization parameter. *Journal of Magnetic Resonance*, 211:178 – 185, 2011.

[4] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Apparent diffusion coefficients from high angular resolution diffusion imaging: estimation and applications. *Magnetic Resonance in Medicine*, 56:395–410, 2006.

[5] M. Descoteaux, R. Deriche, T.R. Knosche, and A. Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transactions on Medical Imaging*, 28:269–286, 2009.

[6] L.R. Frank. Characterization of anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 47:1083–1099, 2002.

[7] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34:561–580, 1992.

[8] D. LeBihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, and M. Laval-Jeantet. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*, 161:401, 1986.

[9] K.E. Sakaie and M.J. Lowe. An objective method for regularization of fiber orientation distributions derived from diffusion-weighted MRI. *NeuroImage*, 34:169–176, 2007.

[10] S.N. Sotiropoulos, D.E. Jones, L. Bai, and T. Kypraios. Exact and analytic Bayesian inference for orientation distribution functions. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2010.

[11] D.S. Tuch, T.G. Reese, M.R. Wiegell, N. Makris, J.W. Belliveau, and V.J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48:577–582, 2002.

Reconstruction, Registration & Multimodality

A Nonlinear Least Squares Method for Solving the Joint Reconstruction and Registration Problem in Digital Breast Tomosynthesis[†]

Guang Yang, John H. Hipwell
{g.yang,j.hipwell}@cs.ucl.ac.uk
David J. Hawkes, Simon R. Arridge
{d.hawkes,simon.arridge}@cs.ucl.ac.uk

Centre for Medical Image Computing
Department of Computer Science
and Medical Physics
University College London
London, UK, WC1E 6BT

Abstract

Digital Breast Tomosynthesis (DBT) offers potential insight into the fine details of normal fibroglandular tissues and abnormal lesions, e.g., masses and micro-calcifications associated with breast cancer, by the production of a pseudo-3D image. In addition, it avoids the superposition, which is usually found in X-ray mammography, with a comparable radiation dose. Algorithms to aid the human observer process DBT data sets involve two key tasks: reconstruction and registration. In established medical image modalities these tasks are normally performed sequentially; the images are reconstructed and then registered. In this paper, we hypothesise that, for DBT in particular, combining the optimisation processes of reconstruction and registration into a single algorithm will offer satisfactory for both tasks. Based on this hypothesis, we have devised a mathematical framework to combine these two tasks, and have implemented both affine and non-linear B-spline registration transformation models as plug-ins. By applying our algorithm to various simulated data, we demonstrate the success of our method in terms of both reconstruction fidelity and in the registration accuracy of the recovered transformations.

1 Introduction

Digital breast tomosynthesis (DBT), is a tomographic modality in which a volumetric image is reconstructed from the acquisition of multiple X-ray images over a limited angular range [3]. By acquiring a 3D image, albeit with coarse depth resolution, DBT aims to disambiguate the overlapping tissues that degrade the sensitivity and the specificity of conventional mammography. In so doing, DBT could be a suitable complementary imaging modality to mammography, enhancing the performance of screening and diagnosis of breast cancer by clinicians.

The workflow in which DBT would be used clinically, involves two key tasks: reconstruction, to generate a 3D image of the breast, and registration, to enable images from different visits to be compared, a task that is routinely performed by radiologists working

[†]Contact Email: G.Yang@cs.ucl.ac.uk. This work has been funded by DTI Project *Digital Breast Tomosynthesis* TP/7/SEN/6/1/M1577G. The authors would like to thank the UK MR Breast Screening Study (MARIBS) for providing the data for this study.

with conventional mammograms. In other modalities, such as MRI and CT, these tasks have traditionally been performed sequentially, i.e. the temporal data sets are first reconstructed independently and then registered. This can be effective if reconstructing using a complete set of data. However, for ill-posed limited-angle problems such as DBT, estimating the deformation is challenging due to the presence of significant reconstruction artefacts, which can lead to severe inaccuracies in the registration.

In this paper, we hypothesise that combining the reconstruction and registration of DBT into a single process will offer satisfactory for both tasks. There is little previous research in this field, and existing techniques applied to modalities other than DBT have focussed on either 2D affine or 3D rigid motion correction rather than 3D affine and non-rigid B-spline implemented in our method. Chung et al. [1] elucidated a combined framework to solve the super-resolution problem of motion correction in MR images, using a 2D affine model. In 2009, Schumacher et al. [2] proposed a method to combine reconstruction and motion correction for SPECT imaging, but only considered 3D rigid motion. The authors have approached the combined problem in DBT using an iterative method before [4, 5].

In addition to the novelty of our fully-coupled simultaneous approach (cf. the iterative method) in incorporating more complex transformation models, we also address the challenging task of applying these techniques to the limited angle datasets acquired in DBT. We test and validate this method using various phantom data, breast MRI, and simulated breast images.

2 Method

Forward Problem: A 3D image, $\mathrm{f}_1^{\mathrm{g}} \in \mathbb{R}^{\mathrm{D}_3}$, two sets of temporal data, p_1, $\mathrm{p}_2 \in \mathbb{R}^{p_{\mathrm{num}} \times \mathrm{D}_2}$, the parametric transformations, $\mathscr{T}_\zeta^{\mathrm{g}}$, and the system matrix, $A \in \mathbb{R}^{p_{\mathrm{num}} \times \mathrm{D}_2 \times \mathrm{D}_3} : \mathbb{R}^{\mathrm{D}_3} \mapsto \mathbb{R}^{\mathrm{D}_2}$, can be related via

$$\mathrm{p}_1 = A\mathrm{f}^{\mathrm{g}} = A\mathrm{R}(\mathrm{x}); \tag{1}$$

$$\mathrm{p}_2 = A\mathscr{T}_\zeta^{\mathrm{g}}\mathrm{f}^{\mathrm{g}} = A\mathrm{T}[\mathscr{T}_\zeta(\mathrm{x})], \tag{2}$$

where p_{num} is the number of limited angle projections, and D_2 and D_3 denote the 3D volume space and 2D projection space. In addition, $\mathrm{f}_1^{\mathrm{g}}$ and $\mathscr{T}_\zeta^{\mathrm{g}}$ are the ground truth of the reconstruction and the parametric transformations respectively, whilst R and T represent the interpolations at original coordinates x and transformed coordinates $\mathscr{T}_\zeta(\mathrm{x})$.

Inverse Problem: We solve the inverse problem by forming the objective function given by

$$\{\mathrm{f}^\star, \zeta^\star\} = \arg\min_{\mathrm{f},\zeta} \left(f(\mathrm{f},\zeta) = \frac{1}{2}\left(\left\|A\mathrm{f} - \mathrm{p}_1\right\|^2 + \left\|A\mathscr{T}_\zeta\mathrm{f} - \mathrm{p}_2\right\|^2\right)\right), \tag{3}$$

in which f denotes the estimation of the unknown volume, and ζ is the estimation of the unknown parametric transformations.

A minimiser $\{\mathrm{f}^\star, \zeta^\star\} \in \mathbb{R}^n$ of $f(\mathrm{f},\zeta)$ is characterised by the necessary condition that the partial derivative with respect to f and ζ equals zero. The partial derivative with respect to f is straightforward, and is given by

$$g(\mathrm{f}) = \frac{\partial f(\mathrm{f},\zeta)}{\partial \mathrm{f}} = A^T(A\mathrm{f} - \mathrm{p}_1) + \mathscr{T}_\zeta^* A^T(A\mathscr{T}_\zeta\mathrm{f} - \mathrm{p}_2), \tag{4}$$

in which $g(\mathrm{f})$ is the gradient with respect to f, and $\mathscr{T}_\zeta^*$ is the adjoint operator of $\mathscr{T}_\zeta$.

To derive the partial derivative with respect to ζ, we apply a small perturbation to the objective function and the linearisation via the norm yields,

$$f\left(\mathrm{f},\zeta+\Delta\zeta\right)=\frac{1}{2}\left(\left\|A\mathrm{f}-\mathrm{p}_1\right\|^2+\left\|A\mathscr{T}_{\zeta+\Delta\zeta}\mathrm{f}-\mathrm{p}_2\right\|^2\right) \tag{5}$$

$$\approx\frac{1}{2}\left(\left\|A\mathrm{f}-\mathrm{p}_1\right\|^2+\left\|A\mathscr{T}_{\zeta}\mathrm{f}+A\frac{\partial\mathscr{T}_{\zeta}}{\partial\zeta}\mathrm{f}\Delta\zeta-\mathrm{p}_2\right\|^2\right). \tag{6}$$

If $g(\zeta)$ denotes the gradient then we have,

$$g(\zeta)=\frac{\partial f(\mathrm{f},\zeta)}{\partial\zeta}=\left(A\frac{\partial\mathscr{T}_{\zeta}}{\partial\zeta}\mathrm{f}\right)^T\left(A\mathscr{T}_{\zeta}\mathrm{f}-\mathrm{p}_2\right)=\left(A\mathscr{T}'_{\zeta}\mathrm{f}\right)^T\left(A\mathscr{T}_{\zeta}\mathrm{f}-\mathrm{p}_2\right).$$

Adjoint Operator of the Transformations: The adjoint operator of the transformation, denoted by $\mathscr{T}^*$, is used to solve the inverse problem. The definition of the adjoint operator in the context of linear transformations on finite dimensional vector spaces is straightforward. By adopting a matrix representation of the linear transformations, we utilise the fact that the adjoint of such a matrix is the same as its transpose.

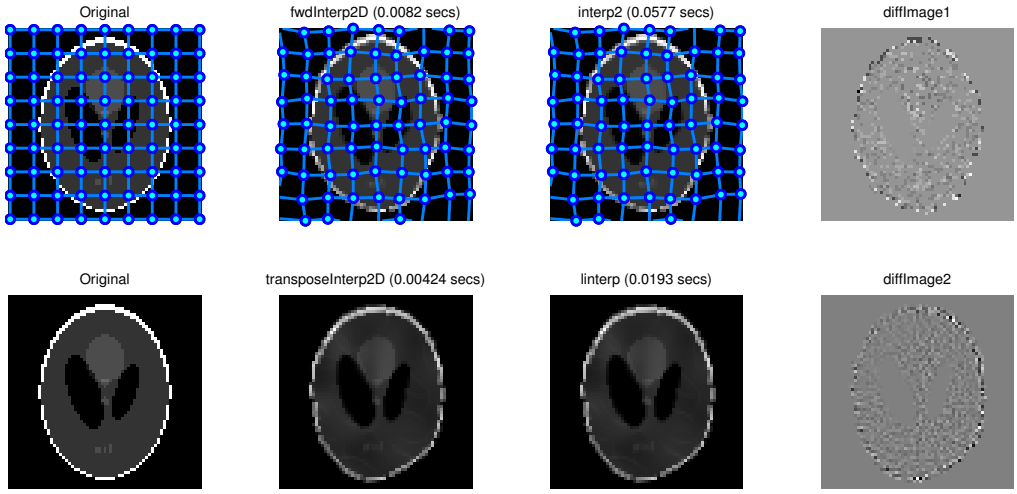


Figure 1: *2D results of the interpolation operation and its transpose. First row: Forward interpolation with deformed grid on; Second row: transpose of the interpolation. From left to right: Original image; Implementation using C with Matlab MEX interfaces; Implementation using Matlab; Difference images between two implementations. The results have shown that our C implementation is faster and accurate.*

The adjoint operator, also known as the Hermitian conjugate, can be defined by

$$\left\langle\mathscr{T}(\mathrm{f}_1),\mathrm{f}_2\right\rangle=\left\langle\mathrm{f}_1,\mathscr{T}^*(\mathrm{f}_2)\right\rangle \tag{7}$$

in which $\langle\cdot,\cdot\rangle$ is the inner product. f_1 and f_2 are arbitrary vectors such that $\forall\mathrm{f}_1,\mathrm{f}_2\in\mathrm{H}_s$, where H_s denotes the Hilbert space; a vector space with an inner product with respect to the associated norm. Although $\mathscr{T}$ is nonlinear with respect to the transformation, ζ, it is linear with respect to the image intensities f. Since $\mathscr{T}_{\zeta}\mathrm{f}=\mathrm{T}[\mathscr{T}_{\zeta}(\mathrm{x})]$,

$$\mathscr{T}_{\zeta}^*\mathrm{f}=\mathscr{T}_{\zeta}^T\mathrm{f}=\mathrm{T}^T[\mathscr{T}_{\zeta}(\mathrm{x})]. \tag{8}$$

In other words, the transpose of an image transformation is the transpose of an interpolation operation. To illustrate this property we implemented both the image interpolation (bilinear interpolation) and the equivalent transpose operation and applied them to 2D (Fig. 1)

and 3D (Fig. 2) test images using randomly created B-spline transformations. In addition, we validated our implementation using Equation 7 with an arbitrary image, to test various transformations. In all cases we obtained the same inner product results.

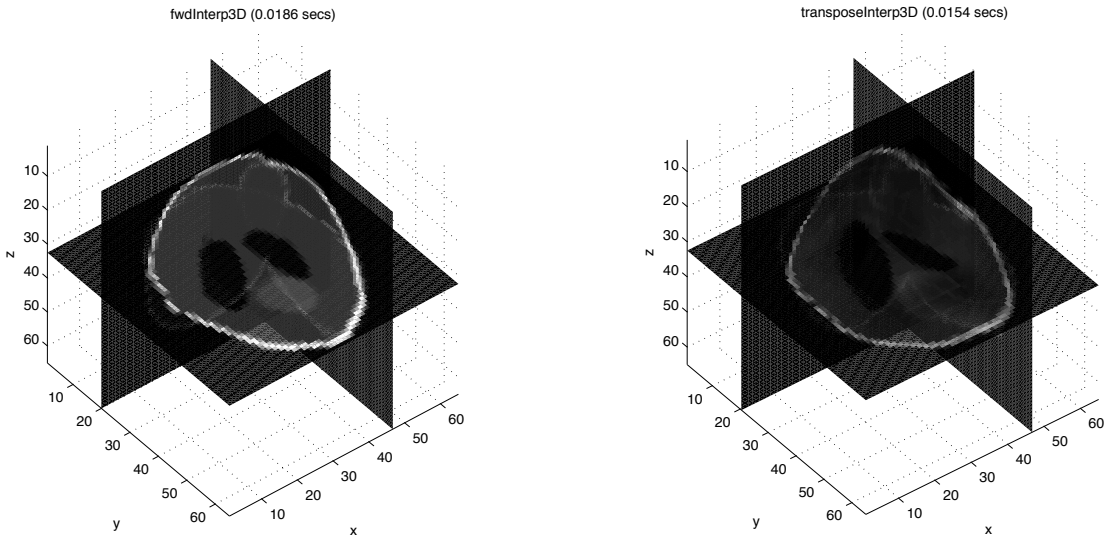


Figure 2: *3D results of the interpolation operation (Left) and its transpose (Right).*

Derivative Operator of the Transformations: The derivative of the transformation operation is a key component of the algorithm and has great impact on the result of the optimisation. Deriving an analytical derivative of the transformation is desirable because it would be fast to compute but is complicated by the need to formulate the derivative of the underlying interpolation. In addition, some interpolation schemes have no analytical derivative. For this reason therefore, we use the Finite Difference Method (FDM) to approximate the derivative operation:

$$\mathscr{T}_{\zeta}^{'} \approx \frac{\mathscr{T}_{\zeta+\varepsilon} + \mathscr{T}_{\zeta-\varepsilon}}{2\varepsilon} \tag{9}$$

where ε is a small number.

Optimisation: The optimisation is performed using a quasi-Newton based Limited Memory BFGS (L-BFGS) method. This approximates the inverse of the Hessian matrix whilst avoiding the considerable memory overhead (for large DBT data sets) associated with computing 2nd order derivatives or their fully dense approximations directly.

3 Results

In this section we investigate the performance of our framework using (a) an affine transformation model and (b) a non-rigid B-spline transformation model.

Affine based experiments: In the first experiment, a 3D toroidal phantom image was created, and subjected to 20 affine transformations to test the robustness of our joint method. In the second experiment, 15 randomly generated affine transformations were applied to a 3D breast MR image. The specific parameters recovered are shown in Fig. 3. In a third experiment, we tested the methods using two MRI acquisitions obtained before and after application of a lateral-to-medial plate compression of the breast. There is no ground truth for the deformation of this dataset, however from the mean squared error (MSE in Table 1), we can conclude that our joint method has successfully reconstructed the data with reasonable registration.

B-spline based experiments: In the fourth experiment, we created a 3D Shepp-Logan phantom image. Although the 3D Shepp-Logan phantom does not represent the structure of the breast, it is a widely used phantom image for tomographic reconstruction. Fig. 4(a) shows central orthogonal slices through the 3D Shepp-Logan phantom ($65 \times 65 \times 65$ in voxels), and the regular grid of B-spline control points for the central slice of the transverse plane. The transformed phantom is shown in Fig. 4(b), with the ground truth transformation. This ground truth deformation is randomly simulated with 9 control points in each dimension using the B-spline transformation model. From the results shown in Figs. 4(c) and (d), we can conclude that our joint method has obtained a reconstruction in high fidelity with an accurate recovery of the non-rigid deformation.

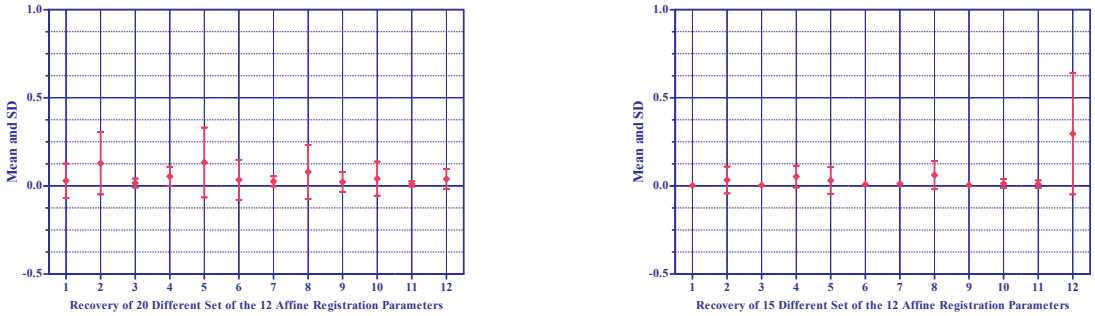


Figure 3: *Mean and standard deviation of the absolute error between the recovered and the ground truth of different sets of affine transformations. Parameters 4, 8, and 12 are the translations along each axis. (Left: Experiment on a 3D toroidal phantom image with 20 randomly created affine transformations; Right: Experiment on a 3D breast MRI image with 15 randomly created affine transformations. The translation could be measured in voxels; however, other parameters have no defined unit because they are calculated using matrices multiplication, e.g., 3D rotation matrix is multiplied by 3D shearing matrix and etc.)*

Table 1: *Comparison of the MSE error* $\frac{1}{\mathbb{N}}\|\mathrm{f}^{\star} - \mathrm{f}^{\mathrm{g}}\|^2$ *before and after performing our joint reconstruction and registration* ($\mathbb{N}$ *is the number of voxels*).

	Initial	**Joint Method**
Toroid Phantom	5.66×10^6	0.24×10^3
Uncompressed Breast MRI	1.18×10^6	3.01×10^3
In vivo DBT simulation	5.32×10^6	3.22×10^4

4 Conclusion

We have presented a method to jointly reconstruct and register temporal DBT datasets and tested it using both affine and B-spline transformation models. Our work has led us to conclude that this joint method produced satisfactory results in both registration accuracy and reconstruction appearance. Furthermore, our framework should be straightforward to incorporate other non-rigid transformation models and priors to regularise the solution. This method has application for the detection of change in temporal DBT data sets. It may also be applied to the combined reconstruction and registration of two view (cranial-caudal (CC) and Mediolateral-oblique (MLO)) DBT data sets, to overcome the null-space limitation of the individual views and produce a single reconstructed volume with improved depth resolution.

References

[1] Chung, J., et al.: Numerical methods for coupled superresolution. Inverse Problems, Vol. 22(4), pp. 1261 (2006)

[2] Schumacher, H., et al.: Combined reconstruction and motion correction in SPECT imaging. Nuclear Science, IEEE Transactions on, Vol. 56(9), pp. 73-80 (2009)

[3] Wu, Tao, et al.: A comparison of reconstruction algorithms for breast tomosynthesis. Medical Physics, Vol. 31(9), pp. 2636-2647 (2004)

[4] Yang, Guang, et al.: Combined reconstruction and registration of digital breast tomosynthesis. IWDM'10, LNCS, Vol. 6136, pp. 760-768 (2010)

[5] Yang, Guang, et al.: Combined reconstruction and registration of digital breast tomosynthesis: Sequential method versus iterative method. MIUA'10, pp. 1-5 (2010)

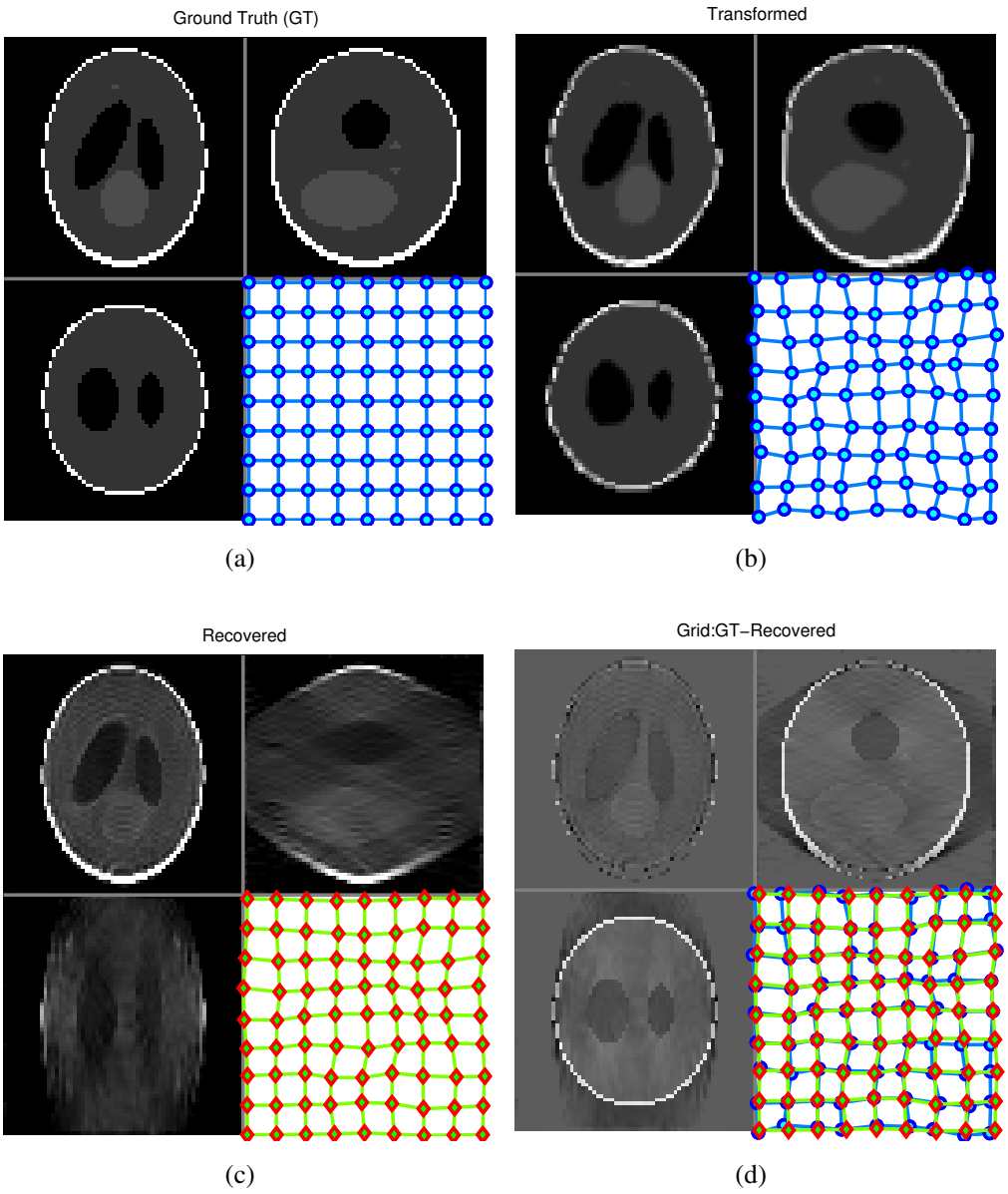


Figure 4: *(a): Original fixed 3D Shepp-Logan phantom and the regular B-spline control point grid for the central slice; (b): Transformed 3D Shepp-Logan phantom and its deformed grid for the central slice, i.e., the ground truth transformation; (c): Joint reconstruction and registration result and its recovered transformation grid for the central slice; (d): Difference image between the joint result and the original fixed image, and the recovered grid superimposed on the ground truth transformation. (Four sub-figures from top to bottom and from left to right are: Transverse view; Coronal view; Sagittal view; Grid of the central slice of the transverse view.)*

Robust Super-Resolution Reconstruction with Multi-Modal Registration and Guidance

Mattias P. Heinrich[1,2]
mattias.heinrich@eng.ox.ac.uk
Mark Jenkinson[2]
Sir Michael Brady[3]
Julia A. Schnabel[1]

[1] Institute of Biomedical Engineering University of Oxford, UK
[2] Oxford University Centre for Functional MRI of the Brain, Oxford, UK
[3] Department of Oncology University of Oxford, UK

Abstract

Increasingly, clinical practice involves acquiring multiple scans with different modalities for diagnostic tasks, as well as longitudinal studies to monitor patient response to therapy and for disease progression. Often, one of these scans is of substantially lower spatial resolution. In order to maintain the highest possible resolution of all scans for subsequent analysis steps, upsampling of the images is needed. Recently, new approaches have been proposed to make use of an inter-modality high-resolution image as a guidance for this upsampling process. While these techniques achieve a significantly higher quality for two perfectly aligned volumes compared to traditional interpolation methods, they are not robust against misalignments between the high- and low-resolution scans. We address this problem by incorporating a deformable multi-modal registration step in the super-resolution reconstruction process. We demonstrate an improved performance of our method compared to two different intra-modality interpolation-based techniques and an inter-modality guided approach without incorporation of registration.

1 Introduction and Background

Medical imaging modalities provide complementary information, due to their different physical acquisition principles. For example, computed tomography (CT) scans have high spatial resolution and good dense tissue contrast, while magnetic resonance images (MRI) excel in higher soft tissue contrast at the cost of a longer acquisition time and lower resolution. The acquisition of scans of multiple modalities is becoming clinical practice. Using their complementary information in the highest possible resolution is therefore of great clinical benefit, for example for pulmonary image analysis of thoracic images, where MRI acquisition is difficult to the breathing motion.

There are two principal challenges involved in the fusion of such multi-modal scans. First, patient motion and imaging distortion, which may have occurred between two scans, has to be compensated for. Rigid and deformable image registration has been widely studied to address this problem. Second, the inherently different spatial resolutions of the different

modalities have to be reconciled to bring both scans into the same coordinate space. Conventionally, interpolation techniques are used to resample the lower resolution (LR) scan to match the higher-resolution (HR) scan (upsampling). However, such an approach uses only the neighbouring intensity values of the same low-resolution scan and ignores the available high-resolution information of the complementary scan; this in turn causes blurring of anatomical structures.

Recently, a number of methods have been proposed which use an inter-modality high-resolution image as a guide for the upsampling process. Kopf et al. [6] introduced a joint bilateral filtering approach for super-resolution (SR) reconstruction based on low-resolution exposure maps, chrominance image and stereo depth maps together with a high-resolution gray-scale image for guidance. In [7] and [8] super-resolution images of low-resolution T2-weighted MRI brain scans are obtained using a non-local weighting process that uses a high-resolution T1-weighted scan as prior. These methods all have in common that the perfect alignment of LR and HR images is required to obtain accurate and robust results. In clinical practice, this assumption generally does not hold true due to residual registration errors, particularly in the case of significantly different voxel dimensions.

We have developed a novel approach, which incorporates the estimation of a deformation field between the multimodal LR and HR scans into the super-resolution reconstruction process. Based on a point-wise multi-modal similarity metric and a diffusion-regularised Gauss-Newton optimisation, a deformation field is computed between the voxels in the LR and HR guidance scans. The deformation field is then used to transform the appropriate non-local weighting of the HR scan in the SR reconstruction. The process is iteratively refined, while the solution is constrained by the underlying imaging physics. Our approach is explained in detail in the next sections. We demonstrate its improved performance on clinically relevant tasks for SR reconstruction, including upsampling of chest MRI images using a HR CT scan as a guide.

2 Guided Upsampling with an Inter-Modality Prior

The observation model of the formation of a LR image $\mathbf{y}$ based on the HR image $\mathbf{x}$ is (as defined in [8]):

$$\mathbf{y} = DBW\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{n}$ represents the noise, D is a sub-sampling matrix, B is a blur operator, and W is the geometric transformation between the HR and LR images. In this work, we assume the noise $\mathbf{n}$ has been removed as a preprocessing step. Alternatively, an appropriate noise-model could be incorporated into the upsampling process. In the case of MRI scanners, the blur operator B can be approximated by a 3D boxcar function.

To ensure that the reconstructed SR volume $\hat{\mathbf{x}}$ complies with the physical acquisition model, an additional constraint, the subsampling consistency (SSC), was introduced by [2]. It requires the downsampled SR image to be equivalent to the original LR image:

$$\mathbf{y} - DBW\hat{\mathbf{x}} = 0 \tag{2}$$

The reconstruction step of our method follows the approach of [7]. Based on the assumption that similar voxels in the inter-modality HR image $\mathbf{z}$ are similar in the reconstructed SR image, we perform a weighted averaging using a joint filtering approach (see e.g. [6]). The weights of this filter are estimated from the non-local neighbourhood Ω of the voxels in

the HR guidance image $\mathbf{z}$; but the averaging is performed on the reconstructed SR voxels. In contrast to linear invariant filters (e.g. Gaussian, Laplacian), the filter characteristic is data-dependent and therefore different for each voxel. Based on a previous estimate $\hat{\mathbf{x}}^t$ the intensities of the SR image can be calculated with the following equation:

$$\hat{\mathbf{x}}^{t+1}(x) = \frac{1}{c} \sum_{r \in \Omega} \hat{\mathbf{x}}^t(x+r) \exp^{-||\mathbf{z}(x)-\mathbf{z}(x+r)||^2/h} \tag{3}$$

where the $\Omega = \{0, \pm 1, \ldots, \pm r_{\max}\}^3$ defines 3D the non-local search region, c is a normalization constant and h is a filtering parameter. A limitation of this approach is the necessity of sufficient contrast in the HR scan. For regions where this is not fulfilled the upsampled anatomical structures will by limited by the lower resolution (see Fig. 2). If the squared intensity distance of the voxels in the HR image $\mathbf{z}$ is replaced by the sum of squared differences (SSD) of small images patches around these voxels, Eq. 3 can be seen as joint version of the popular non-local means filter [3]. In [7], the filter parameter h is chosen empirically and is reduced by a factor of 2 after each iteration, thus introducing a coarse-to-fine scheme of the SR reconstruction. The reconstruction process is performed iteratively by alternating steps of guided reconstruction (Eq. 3) and enforcement of the SSC (Eq. 2).

3 Robust SR using Multi-Modal Registration

In previous work [7], [8] the geometric transformation W in Eq. 1 was assumed to be known a priori. However, as shown in [7], even slight initial misalignments cause the reconstruction performance to deteriorate. The reconstructed SR volume is at best as good as an interpolated version of the LR image (without using the HR guidance). In some cases, it can also cause artifacts in the reconstructed SR image, as shown in Fig. 2 (c).

We address this problem by incorporating deformable multi-modal registration between the LR volume $\mathbf{x}$ and the HR volume $\mathbf{z}$. In [5] a point-wise multi-modal similarity metric is introduced, which can cope with the complex nature of multi-modal similarity. A multi-dimensional descriptor is computed at each location in both images, which is modality-independent and discriminative to prominent image features (such as edges and corners). We follow a similar approach as [5] and optimise a cost function consisting of the voxel-wise L2 norm of the multi-dimensional descriptors and a diffusion regularisation in a multi-resolution Gauss-Newton framework. To reduce the influence of smoothing due to interpolation, we apply the resulting deformation to the guidance image and recalculate the weights in Eq. 3 based on the transformed image at each iteration.

When initialised with SR=LR, our proposed approach consists of three alternating steps:

1. multimodal deformable registration of SR and HR scans to find W
2. reconstruction based on guided filtering using the inter-modality HR scan ($r_{\max} = 3$)
3. intensity correction to ensure subsampling consistency (SSC).

4 Experiments and Results

We demonstrate our proposed method on three different datasets. First, we use a simulated 3D MRI T1/T2 brain phantom (Brainweb [4]), which has also been used in [7] and [8]

and allows us to compare our method with published results. Second, we use the Visible Human Dataset [1] consisting of 3D MRI chest scans with T1 and T2 weighting, which have been acquired post-mortem and therefore do not have any significant misalignement. For these two datasets, we artificially subsample one of the two scans to obtain a low-resolution volume. Third, we use the same chest MRI scan together with an additional CT scan of the same subject, where there exists a residual non-rigid mis-alignement between the two scans. The reconstructed SR volume $\hat{\mathbf{x}}$ is compared with the known ground truth volume $\mathbf{x}$, and the reconstruction error can be defined as mean squared error (MSE) or peak to noise ratio (PSNR) in dB: $MSE = \frac{1}{n}\sum(\hat{\mathbf{x}}-\mathbf{x})^2$ and $PSNR = 10\log_{10}\frac{(\max\hat{\mathbf{x}})^2}{MSE}$.

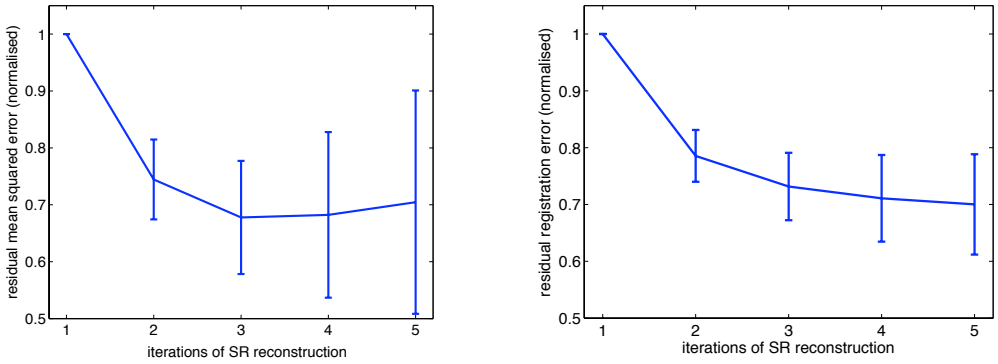


Figure 1: Residual error after each iteration of our proposed SR reconstruction method applied to the Brainweb dataset and averaged over different subsampling factors (2, 3, 4 and 6 mm). Mean squared error (MSE) is displayed on the left, the target registration error (TRE) on the right. Results are normalised by the value of the first iteration.

Table 1: Performance of reconstruction / interpolation methods (PSNR in dB) for noise-free Brainweb (BW) volumes and chest scans of the Visible Human Dataset (VHD). The BW HR guidance image (T1-weighted MRI) has isotropic voxel-size of 1.0 mm^3. The VHD HR scan (MRI-T1) has a voxel-size of 1.9x4.0x1.9 mm^3. A random subpixel shift of up to 3 mm has been applied to the HR scan before reconstruction. The LR scan is a T2-weighted MRI.

Image dataset	subsampling	NN	B-spline	**proposed**	w/o registration
Brainweb	2 mm	27.9 dB	31.5 dB	**33.7 dB**	28.2 dB
	3 mm	24.0 dB	26.0 dB	**30.5 dB**	25.0 dB
	4 mm	22.5 dB	23.7 dB	**28.1 dB**	23.6 dB
	6 mm	20.5 dB	21.2 dB	**23.9 dB**	21.4 dB
Visble Human	2 mm	31.6 dB	33.6 dB	**33.6 dB**	29.7 dB
	3 mm	28.7 dB	30.0 dB	**30.8 dB**	26.9 dB
	4 mm	27.2 dB	28.3 dB	**29.6 dB**	25.6 dB
	6 mm	25.5 dB	26.4 dB	**27.7 dB**	24.0 dB

We perform our proposed SR reconstruction on LR volumes with varying subsampling factor (2, 3, 4 and 6 mm). A random (subpixel) translation of up to 3 mm is applied for the first and second experiment to simulate initial misalignments. The resulting MSE between reconstructed SR and the ground truth volume and the residual registration error after each iteration are visualised for the first experiment in Fig. 1. To enable comparison between different subsampling factors the results are normalised by the first value of the first iteration.

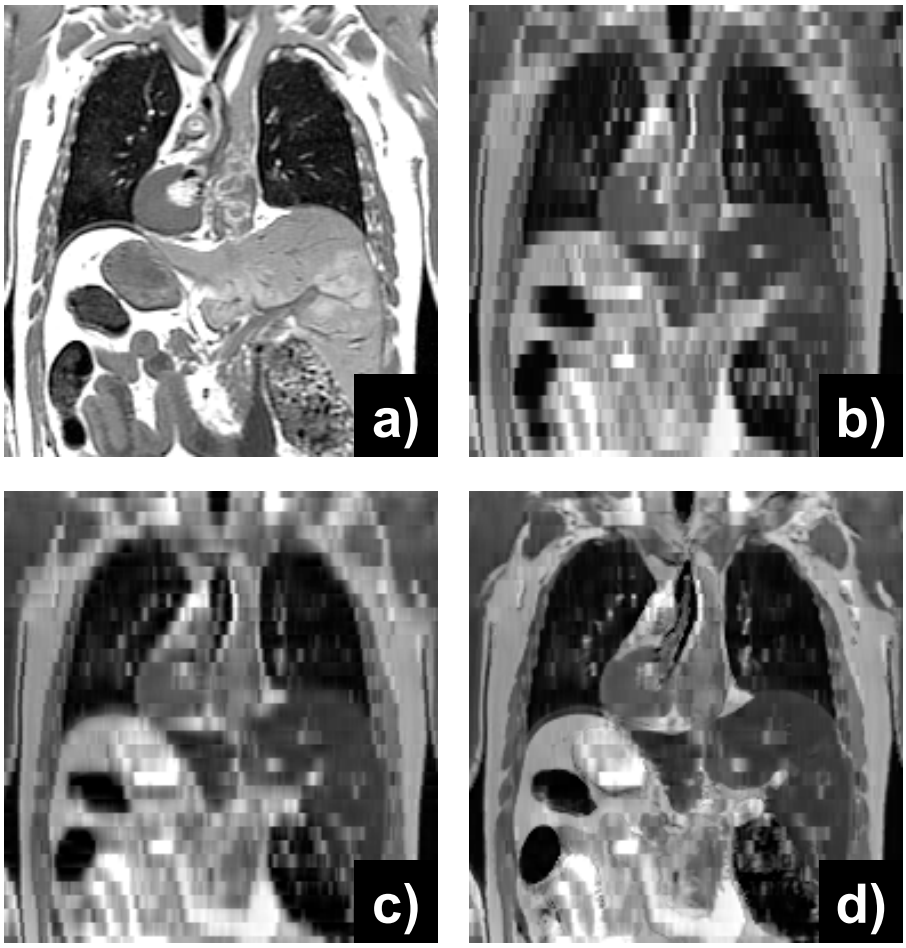


Figure 2: Coronal plane of chest MRI of VHD. (a) HR T1-weighted guidance volume. (b) LR T2-weighted volume (subsampled by a factor of 6). (c) SR reconstruction without registration. (d) SR reconstruction of our proposed method yields are visually clearly improved result. Problematic are areas, which have no corresponding anatomy in the HR scan.

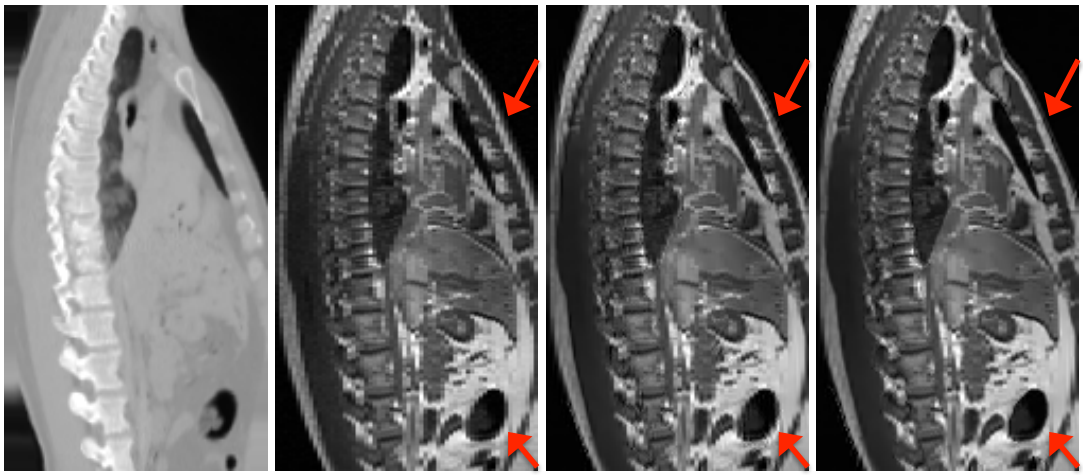


Figure 3: Sagittal plane of chest CT/MRI scan pair. From left to right: HR CT guidance volume, LR T1-weighted volume (NN-interpolation), SR reconstruction without registration, SR reconstruction of our proposed method. The red arrows indicate areas in where our proposed method achieves an improved SR reconstruction (reduced aliasing along edges).

For both measures the initial error could be further reduced by 30% on average using our proposed technique. An overview of the results for the first two experiments is given in Table 1 and an example outcome of the SR reconstruction of a chest MRI upsampled by a factor of 6 is shown in Fig. 2.

For the third experiment, no ground truth is available, therefore only visual results can be presented. The MRI scan (T1-weighted) has a resolution of 1.875x4.0x1.875 mm, the HR CT reference of 0.9x0.9x1.0 mm. We use the proposed method to construct a super-resolution MRI volume with a voxel-size of 1.875x2.0x1.875 mm. The scans are manually rigidly aligned, but there is still some minor non-rigid misalignment, due to slightly changed body positioning and geometric distortion of the scanners. Figure 3 shows an improved reconstructed volume compared to the nearest neighbour (NN) interpolation and the guided SR reconstruction without reconstruction (following the approach of [7]).

5 Conclusion

We have presented a new method to reconstruct a super-resolution (SR) volume based on a low-resolution scan and an inter-modality high-resolution guidance volume. Our technique incorporates a multi-modal registration step for a robust SR reconstruction. We demonstrate on a number of challenging clinical datasets, that the registration accuracy and the SR reconstruction can be improved by combining both methods into a unified framework. One remaining challenge are areas in the LR, which have no corresponding anatomy in the HR scan. In the future, we would like formulate our approach in a combined energy functional (registration and SR reconstruction) and apply it on a larger set of clinical CT/MRI scans.

References

[1] M.J. Ackerman. The visible human project. *Proc. of the IEEE*, 86(3):504 –511, 1998.

[2] J. Banerjee and C.V. Jawahar. Super-resolution of text images using edge-directed tangent field. In *Document Analysis Systems, 2008. DAS 2008.*, pages 76 –83, sept. 2008.

[3] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *CVPR 2005*, volume 2, pages 60–65, 2005.

[4] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, and A.C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Transaction on Medical Imaging*, 17(3):463–468, 1998.

[5] M.P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. Gleeson, M. Brady, and J.A. Schnabel. Non-local shape descriptor: A new similarity metric for deformable multi-modal registration. In *MICCAI 2011*, volume 6892 of *LNCS*, pages 541–548. 2011.

[6] J. Kopf, M.F. Cohen, D. Lischinski, and M. Uyttendaele. Joint bilateral upsampling. *ACM Transactions on Graphics (SIGGRAPH 2007)*, 26(3), 2007.

[7] J. Manjon, P. Coupe, A. Buades, D.L. Collins, and M. Robles. MRI superresolution using self-similarity and image priors. *Intl. J. Biomed. Imag.*, 2010:1 – 11, 2010.

[8] F. Rousseau. A non-local approach for image super-resolution using intermodality priors. *Medical Image Analysis*, 14(4):594 – 605, 2010.

Multimodality investigation of microstructure: the combination of diffusion NMR and diffuse optical spectroscopy

Alessandro Proverbio[1]
alessandro.proverbio.09@ucl.ac.uk
Dr Bernard Siow[23]
Dr Mark Lythgoe[2]
Prof. Daniel C. Alexander[34]
Dr Adam P. Gibson[1]

[1] Department Medical Physics and Bioengineering, UCL, London, UK
[2] Centre for Advanced Biomedical Imaging, UCL, London, UK
[3] Centre for Medical Image Computing, UCL, London, UK
[4] Department Computer Science, UCL, London, UK

Abstract

We propose that multimodality imaging may be enhanced by combining datasets with a model of the underlying physics prior to image reconstruction, as opposed to the traditional approach of combining the reconstructed image. Here, we demonstrate this approach by merging near infrared diffuse optical spectroscopy with diffusion NMR to inform a model describing the microstructure of a sample. A computational validation of the results, produced with synthetic datasets, shows the enhanced accuracy of the estimation, while an application of the model on experimental measurements confirms the validity of the models.

1 Introduction

Traditional multimodality imaging relies on combining reconstructed images to produce information which is not available when using one modality alone. Here, we propose a new approach which is to use data acquired prior to reconstruction to inform a model of the underlying physics of the phenomenon under examination.

This work provides a proof of principle of the idea that the fusion of two modalities increases the accuracy compared to using a single modality. A multimodality investigation of microstructure has been designed by combining near infrared diffuse optical spectroscopy (DOS) and diffusion NMR (dNMR). Ultimately, we intend that this approach will lead to a new method for imaging tumour microstructure.

The modalities inform a common model able to estimate the microstructure parameters of interest. The two modalities measure different aspects of the physics. DOS is sensitive

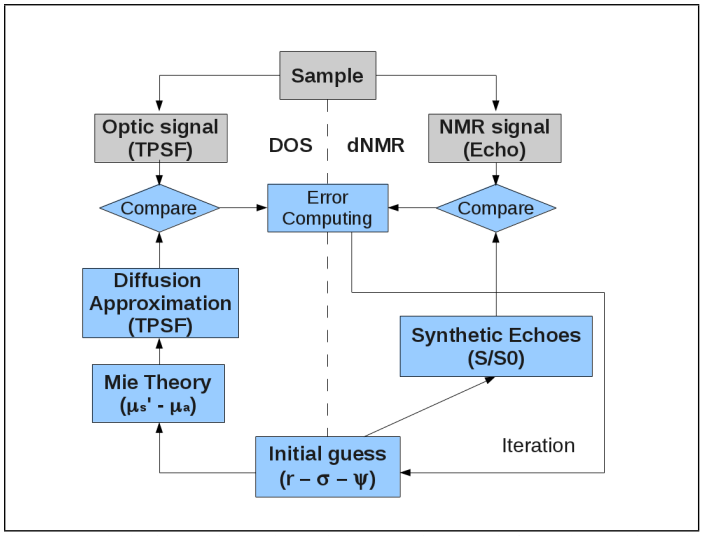


Figure 1: A scheme explaining the algorithm adopted for the mixed model approach.

to the absorption, and to the number and the size of scatterers in the sample, while dNMR observes the effect of molecular diffusion constrained by microstructural compartments.

2 Methods

We derive the distribution of size and the volume fraction of oil particles suspended in a water-based compound. The sample was commercial light mayonnaise (J Sainsbury plc, Basic Mayonnaise). It is a homogeneous and isotropic material with a stable microstructure composed of microspheres in a dispersion medium sensitive to DOS and to dNMR. The optical signal is affected by the change of refractive index between oil globules and dispersion medium, while the dNMR depends on the constrained diffusion of oil particles. The dispersion medium is predominantly water with some other components (proteins, carbohydrates) and within it are oil droplets. The microstructural parameters of interest are the distribution of radii of the microspheres, modeled as log-Normal, with mean r, standard deviation σ, and volume fraction ψ.

2.1 NIR diffuse optical spectroscopy

Time-domain DOS consists of the observation of Temporal Point Spread Functions (TPSF) recorded in transmission through a homogeneous slab when the number of scattering events is large enough for the light transport to be assumed to be diffusive.

We used time-domain DOS to build up a histogram of the time taken for individual photons to diffuse through a homogeneous slab of mayonnaise. The measurements were obtained using a time-domain optical tomography system known as MONSTIR [4]. The sample was contained in a $17 \times 48 \times 52 mm^3$ transparent plastic box and the source was a laser that emitted 5ps pulses of light at 780nm. On the opposite face, a $5.5mm$ diameter detector recorded the TPSF. The measurements were averaged for a duration of $10s$ and the final result was calibrated by deconvolving it with a reference measurement acquired without the sample.

The TPSF is fitted with a Diffusion Approximation (DA) model [1], using a Levenberg-Marquardt algorithm. The apparent scattering μ_s' and absorption μ_a coefficients are esti-

mated. From the optical parameters, a linear interpolation of the dispersion medium and oil absorption coefficients are able to provide the volume fraction and the refractive index as

$$\begin{aligned} n_s &= \psi n_i + (1-\psi) n_o \\ \mu_{a,s} &= \psi \mu_{a,i} + (1-\psi) \mu_{a,o} \end{aligned} \tag{1}$$

where n is the refractive index (the subscripts o indicates the dispersion medium compartment, i the microspheres, and s the sample as a whole). The scattering coefficient is provided by the Mie theory [6], depending on the ratio $n_r = n_i/n_o$ and the shape factor $k = r/\lambda$, where λ is the wavelength of the incident beam. Mie theory provides the scattering coefficient $\mu_{s,Mie}$ and anisotropy factor g_{Mie} relative to a single scattering event, and the scaling on the number of particles per unit of volume provide the sample apparent scattering coefficient as follows:

$$N_V = \frac{3\psi}{4\pi r^3}, \qquad \mu'_{s,s} = N_V \, \mu'_{s,Mie}. \tag{2}$$

The use of the optical modality alone is not able to define the distribution of radii σ but only the mean value r. However, optics can determine the volume fraction ψ.

2.2 Diffusion NMR

For Diffusion NMR modality, we applied a Stimulated Echo sequence to a $20 \times 20 \times 20 mm^3$ sample using a 9.4T Varian preclinical scanner. The signal attenuations for the different parameters were calculated by normalizing the peak amplitudes in the diffusion weighted spectra with the corresponding peaks the unweighted spectra. These attenuations represent the effect of constrained diffusion of oil and water molecules. A set of measurements with a combination of diffusion times $\Delta = 50,\ 100,\ 150,\ 200,\ 250,\ 400,\ 700ms$ and gradients $G = 0,\ 0.1,\ 0.2,\ 0.4,\ 0.6,\ 0.8,\ 0.95T/m$, with $TR = 4s$ and $\delta = 5ms$ was obtained.

The signal attenuations were fitted with the analytical model representing the diffusion in spherical compartments as presented in literature [3]. Since the size of microspheres contained in the sample may be described by a log-Normal distribution of probability, the signals produced by all the possible radii have been integrated to obtain the global contribution [2]. The fitting of the the experimental data with the model discussed has been implemented with a Levenberg-Marquardt algorithm, and the diffusivity coefficient D has also been estimated. The dNMR modality can provide r and σ but not ψ.

2.3 Mixed model

The multimodality approach fits the microstructural parameters by optimizing the difference between the experimental and the analytical signals in the two modalities as shown in figure 1. The difference between the experimental signal and the simulated value in each modality are combined as

$$E_{MIX} = w \, ||TPSF_{EXP} - TPSF(r, \psi, t_0)|| + (1-w) \, ||S_{EXP} - S(r, \sigma, D)|| \tag{3}$$

where w is the weighting factor, set to 0.5. In this formula, the subscription EXP refers to data measured during the experiments while the others are generated by the analytical model. $TPSF$ describes the signal in the DOS modality, while S the echoes attenuation. E_{MIX} is minimized and an optimal solution is retrieved. Since the optical modality is unable to estimate the distribution of radii, but only their mean value, a representation of radii

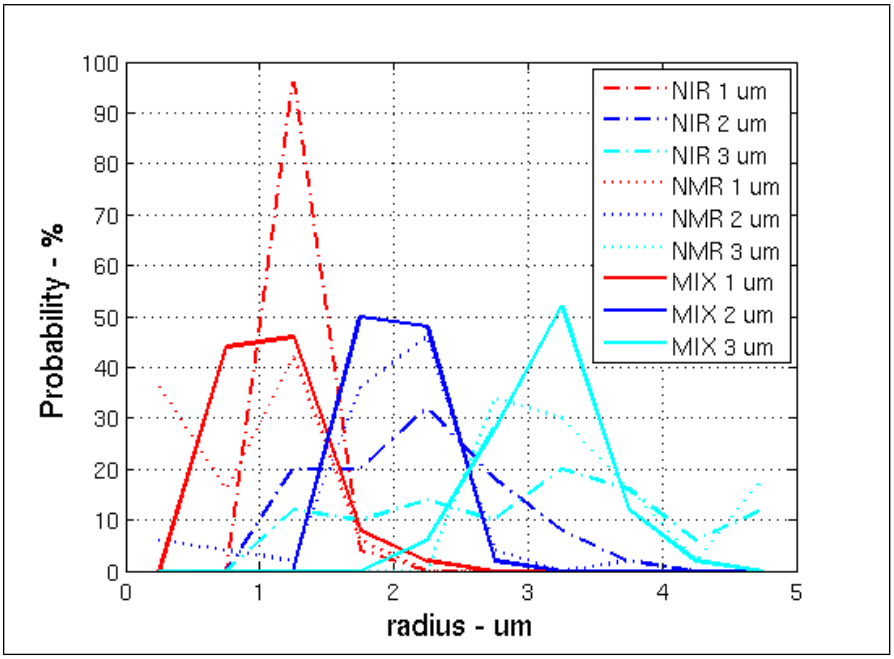


Figure 2: Fitting of synthetic dataset generated with different values of r and SNR = 16. The dataset are fitted with the two modalities independently and with the combined approach, and the estimated values of r are shown in this graph.

distribution has been adopted to match the information contained in dNMR data to make them compatible with DOS modality. The optical properties are generated by considering a poly-disperse distribution of microsphere sizes [5].

3 Results

In order to validate the model, 50 synthetic datasets were generated for every set of microstructural parameters. Noise was added such that the SNR was 16, and the parameters considered were $r = 1,\ 2,\ 3\,\mu m$, $\sigma = 1,\ 2,\ 3\,\mu m$ and $\psi = 0.1,\ 0.2,\ 0.3$. Here, on every single dataset, the three models have been applied and the estimated values of average radius are shown in figure 2. Table 1 shows the mean and the associated error of the fitted values of r averaged over the 50 dataset. The combined model produces more accurate results in terms of average r fitted and associated error, and the improvement against the single modalities increases with the size considered. DOS tends to provide better estimations than dNMR in terms of fitted values and error of estimation, but figure 2 shows some bad estimations for lower values of r which are not highlighted by the data in table 1. Further, an increased number of repetitions in the NMR scans may improve the performances of the method with a resulting increase of acquisition time.

The estimation of the parameters σ and ψ follow the same trend, but the results are not explicitly shown. The quantification of σ, which directly depends only on dNMR, is improved indirectly by the estimation of r using DOS in this combined approach. In the same way, ψ drastically increases its accuracy in the mixed model especially for the ψ of interest (around 30%).

The fitting of experimental data is shown in figure 3. DOS is able to fit the TPSF when the sample is sufficiently thick as shown in figure 3(a). The number of scattering events is

model	$1\mu m$	$2\mu m$	$3\mu m$
DOS	1.09±0.16	2.13±0.67	3.06±1.34
dNMR	0.82±0.57	1.93±0.63	3.75±1.49
MIX	1.02±0.31	1.98±0.21	3.14±0.38

Table 1: Average estimated value of r and error of the data shown in figure 2.

sufficiently high to ensure the applicability of the diffusion approximation. The good fitting is not related to the application of Mie theory but only on an analytical description of the signal. dNMR signal presents a good fitting shown in 3(b). The fitted value of diffusivity coefficient is $D = 0.0098(\mu m)^2/ms$ which is in agreement with the expected values found in literature [2].

4 Discussion

The use of a single modality provides a partial solution of the problem. DOS estimates r and ψ, while dNMR provides r and σ but not ψ. Only their combination may obtain the whole set of parameters at the same time.

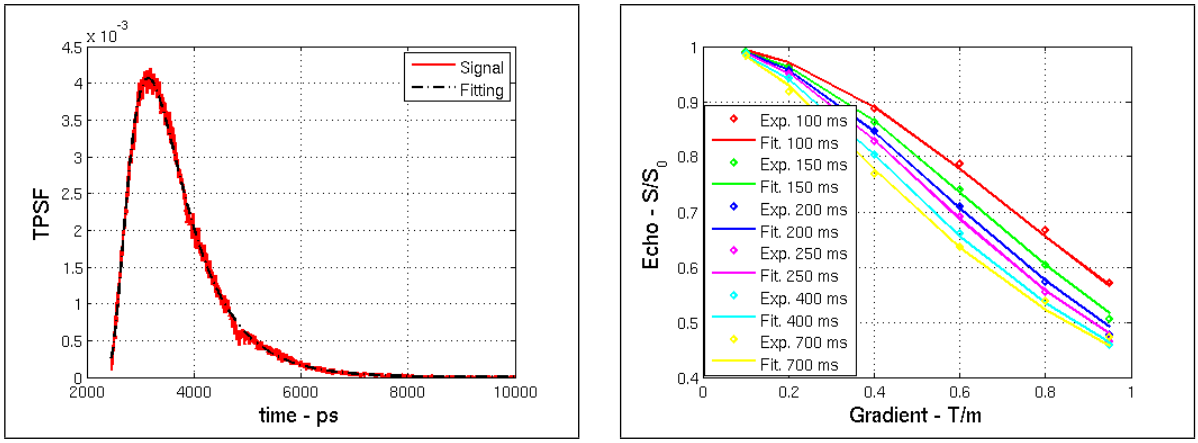
The DOS model requires knowledge of the absorption coefficient and refractive index of the two compartments, which have been retrieved from the literature. The parameters of the inner compartment are well known while the absorption of the dispersion medium has been deduced from its chemical composition. Further experiments are required to obtain the absorption coefficients and refractive indexes more accurately. A second point is the consideration that Mie theory identifies similar pairs (μ_s', μ_a) for different microstructural parameters; therefore, a precise choice of the initial guess is required in order to converge to an accurate solution. This dependency on initial guess is the main disadvantage of this modality. A further disadvantage is the lack of sensitivity to the distribution of sizes.

The dNMR model is instead able to fit the echoes very well. No input parameters are required, and the obtained value of the diffusivity coefficient is a confirmation of the accuracy of the model, which appears close to the expected value from literature. The model is very sensitive to the inner microstructure and the fitting depends only slightly on the initial guess. The time required for the fitting is short and the model is well tested. This modality is limited by the fact that it does not provide information about the volume fraction, but the model fits the real data accurately.

The combined model provides an enhanced accuracy with respect to the single modalities. It is insensitive to the initial guess since it exploits the dNMR properties, but the fitting of distribution width is improved thanks to the contribution of optical modality which cross-informs the model about r. Finally, the estimation of concentration is also improved due to the contemporary fitting of the size with a robust model such that of dNMR.

5 Conclusion

The use of combined models increases the accuracy of the fitting because of the cross-talking between the single techniques. The approach is general and may be extended to other modalities without any complex additional work. The necessity of a more advanced validation, such as a microscopy investigation, is strong. Finally a calibration of the input parameters and of the weight w is required.



(a) (b)

Figure 3: Fitting of experimental data obtained from mayonnaise in DOS modality (a), and in the dNMR modality (b).

The extension of the microstructural model to anisotropic cases will provide practical applications to biological tissues, and the ultimate result will be the solution of an imaging problem with the multimodality information of a microstructural model in a three dimensional space.

6 Acknowledgments

The authors would like to thank Louise Enfield, the Centre for Advanced Biomedical Imaging, and EPSRC who funded this work (grants EP/F01208X/1, EP/E007748 and EP/H046410/01).

References

[1] D. Contini, F. Martelli, and G. Zaccanti. Photon migration through a turbid slab described by a model based on diffusion approximation. I. Theory. *Applied Optics*, 36 (19):4587–4599, 1997.

[2] P. Denkova and et al. Evaluation of the precision of drop-size determination in oil/water emulsions bi low-resolution NMR spectroscopy. *American Chemical Society*, (20): 11402–11413, 2004.

[3] J.S. Murday and R.M. Cotts. Self-diffusion coefficient of liquid lithium. *The Journal of Chemical Physics*, 48(11):4938–4945, 1968.

[4] F.E. Schmidt and et al. A 32-channel time-resolved instrument for medical optical tomography. *Review of Scientific Instruments*, 71(1):256–265, 2000.

[5] H.J. van Staveren and et al. Light scattering in Intralipid-10% in the wavelength range of 400–1100nm. *Applied Optics*, 30(31):4507–4514, 1991.

[6] W. Wiscombe. Improved Mie scattering algorithms. *Applied Optics*, 19(9):1505–1509, 1980.

Ensemble Learning Incorporating Uncertain Registration

Ivor J.A. Simpson[1,2]
ivor.simpson@eng.ox.ac.uk
Julia A. Schnabel[1]
julia.schnabel@eng.ox.ac.uk
Jesper L.R. Andersson[2]
jesper@fmrib.ox.ac.uk
Adrian R. Groves[2]
adriang@fmrib.ox.ac.uk
Mark W. Woolrich[2,3]
mark.woolrich@ohba.ox.ac.uk

[1] Institute of Biomedical Engineering
Old Road Campus Research Building
Oxford OX3 7DQ

[2] FMRIB Centre
John Radcliffe Hospital
Oxford, OX39DU

[3] Oxford Centre for Human Brain Activity
Warneford Hospital
Oxford OX3 7JX

Abstract

In this paper we propose a novel approach for improving the robustness of quantitative prediction methods in spatially normalised statistical analysis of magnetic resonance (MR) images of the human brain. This is achieved by estimating the distribution of image registration mappings from subject to atlas space, rather than just using the *maximum-a-posteriori estimate*. As any derived predictions are highly dependent on the registration, the distribution of spatially normalised feature data can be derived from the set of probable mappings. This distribution of feature data can then be used to generate alternative training examples to create multiple predictors, which can then be combined in an ensemble learning approach. This allows for a more generalised prediction, which compensates for the inherent uncertainty in registration. Furthermore, extra testing examples can be generated to provide a measure of the prediction uncertainty. We demonstrate that using an ensemble learning approach with random feature data samples always leads to improvement in classification rate when separating subjects with Alzheimer's Disease from normal controls using a linear support vector machine.

1 Introduction

Medical imaging data is often used to make quantitative predictions about the current or future state of a subject. Machine learning techniques such as statistical classifiers and regressors are often used to facilitate this objective. Most of these machine learning techniques are *supervised*, which means they require a training set of data from which the methods learn about the relationship between the feature data $\mathbf{d}$, e.g. voxel intensities, and outcome variable o, e.g. disease score/group. When presented with unseen feature data, these methods should be able to predict the true value with a high degree of accuracy.

Most standard machine learning algorithms require feature data of all subjects to be transformed into a common frame of reference to allow comparison. To map to this reference space, a registration tool is used to estimate the mapping between the subject image and atlas space (usually an average or representative subject). From this mapping, subject feature data can be transformed to the atlas space, a process known as spatial normalisation. This spatially normalised data can be then be examined in a voxel based morphometry (VBM) framework. Additionally, features of interest can be derived from the estimated transformation, a process known as deformation tensor based morphometry (TBM).

As has been previously shown, inter-subject brain registration is not an exact process [4]. Therefore the spatially normalised features, which are used as a basis for making predictions, are highly dependent on the registration procedure and are unlikely to be perfectly aligned. This misalignment of data will be a contributing factor to any incorrect predictions. The majority of registration methods simply estimate the *maximum-a-posteriori* (MAP), which is the most likely mapping subject to some constraints. However, recent registration methods have emerged which allow estimates of the registration uncertainty [6][8]. This facilitates the consideration of the set of probable mappings, as opposed to just the MAP.

We take an ensemble learning approach [2], where multiple statistical predictors are averaged. To create multiple classifiers we use a parametric variant of bootstrapping [3] to produce new examples of all the training subjects according to the distribution of the feature data. For spatially normalised feature data, the distribution of the feature data can be estimated based on the set of probable mapping inferred from the registration algorithm. Therefore, in this work, we randomly draw samples of feature data to train statistical predictors. This is repeated to generate an ensemble of predictive models. This ensemble accounts for the inherent registration uncertainty. Samples of the data distribution for test subjects also allows the estimation of the variability in prediction under uncertain registration.

In this work we propose an ensemble learning scheme using a parametric bootstrap approach to provide robust classification which accounts for the registration uncertainty. We can apply this approach to voxel, or deformation tensor based morphometry approaches to estimate differences between subject groups using any standard black box prediction tool. To demonstrate the effect of this method, we apply it to discriminating between subjects with Alzheimer's disease and age matched healthy controls using a standard black box classifier. In our experiments we find that using ensemble learning with data features sampled according to registration uncertainty leads to a consistent improvement in classification accuracy.

2 Methods

2.1 Probabilistic Registration Methods

A probabilistic registration method that can estimate posterior transformation distributions is required to estimate the distribution of the feature data, which accounts for the uncertainty in registration. Standard registration procedures use a MAP approach to infer the mapping between images. These approaches do not provide any estimates on the confidence of the inferred mapping, and consequently do not lend themselves to this work.

There are two published methods which we are aware of which allow the inference of a distribution of probable mappings: Risholm et al. [6] use MCMC to numerically estimate the full posterior distribution of transformation parameters whilst marginalising over the regularisation parameters. This method is computationally very expensive even for low degrees

of freedom. The alternative approach proposed by Simpson et al. [8] uses Variational Bayes (VB) to infer an approximate posterior distribution of the set of transformation, regularisation, and noise parameters. Although this method is more efficient, it uses the assumption that the true distribution of transformation parameters follows a multi-variate normal distribution, which so far has not been justified as necessarily being appropriate. Nevertheless, we choose to use the method of Simpson et al. due to the large computational benefits. We include a brief summary of their method here:

Image registration can be described using a generative model as given in equation

$$\mathbf{y} = \mathbf{t}(\mathbf{x}, \mathbf{w}) + \mathbf{e} \tag{1}$$

where $\mathbf{y}$ is the target image, $\mathbf{t}(\mathbf{x}, \mathbf{w})$ is the transformed source image $\mathbf{x}$ and $\mathbf{w}$ parameterises the transformation. $\mathbf{e}$ models the image mismatch where $\mathbf{e} \sim \mathcal{N}(0, \phi^{-1}\mathbf{I})$, $\mathbf{I}$ is the matrix identity and ϕ is the global inverse variance. We use a Free-Form Deformation (FFD) transformation model [7], where $\mathbf{w}$ is the set of b-spline knot displacements. Priors are included on all our unknown model parameters. Most importantly, a prior on $\mathbf{w}$ is required to provide regularisation of the mapping, $P(\mathbf{w}) = \mathcal{N}(0; (\lambda\Lambda)^{-1})$ where Λ encodes the bending energy regularisation, and λ is an inferred parameter which controls the strength of the prior.

Posterior distributions are inferred using a set of iterative update equations applied until parameter convergence. Of particular interest is the posterior distribution of $\mathbf{w}$, $P(\mathbf{w}|\mathbf{y}) \sim q(\mathbf{w}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}^{-1})$. The update equations for the hyper-parameters are given as:

$$\boldsymbol{\Upsilon} = (\alpha\bar{\phi}\mathbf{J}^T\mathbf{J} + \bar{\lambda}\boldsymbol{\Lambda}) \tag{2}$$

$$\boldsymbol{\Upsilon}\boldsymbol{\mu}_{new} = \alpha\bar{\phi}\mathbf{J}^T(\mathbf{J}\boldsymbol{\mu}_{old} + (\mathbf{y} - \mathbf{t}(\mathbf{x}, \boldsymbol{\mu}_{old}))) \tag{3}$$

where $\boldsymbol{\mu}_{old}$ is the previous mean estimate of the transformation parameters $\mathbf{w}$ and $\mathbf{J}$ is the matrix of first order partial derivatives of the transformation parameters with respect to $\mathbf{t}(\mathbf{x}, \boldsymbol{\mu}_{old})$. $\bar{\lambda}$ and $\bar{\phi}$ are the expectation of the regularisation and image noise distributions, respectively. α is a factor which models the correlation in the image mismatch.

$q(\mathbf{w})$ is the approximate posterior distribution of the inferred transformation parameters. The shape and scale of this distribution is dependent on the structure of the **image information**, weighted by the **noise precision**, which indicates the level of image mismatch. It also depends on the form of the **spatial prior**, e.g. the bending energy, weighted by the **spatial precision** which is related to the similarity of the transformation to the spatial prior.

2.2 Using sampled mappings

Once the registration model parameters have been inferred, we can consider taking a voxel- or deformation tensor-based morphometry [1] approach to provide a framework for the classification of subjects into their respective groups.

In the standard setting, VBM is performed by transforming the subject data to the atlas space based on the expectation of the transformation distribution, $\mathbf{d} = \mathbf{t}(\mathbf{x}, \boldsymbol{\mu})$. In TBM, instead of examining the spatially normalised image information, the assumption is made that the discriminative differences between subjects are contained in the deformation field used to map each subject to the template image. Feature data is often derived from the voxelwise 3×3 Jacobian matrix of the transformation of the mapping $\mathbf{J}_m$. Often a scalar measure of the Jacobian matrix is used for comparison, most commonly $\mathbf{d}_i = \log|\mathbf{J}_m|_i$, where i is a voxel index. $\log|\mathbf{J}_m|_i$ provides a measure of the expansion/contraction of a particular voxel as a

result of the mapping. The spatially normalised images can be compared between images in either a voxelwise or a multivariate fashion as features for classification. By sampling from the approximate posterior distribution of mappings $q(\mathbf{w})$ an estimation of the distribution $P(\mathbf{d}|\Theta)$, where $\Theta = \{\mathbf{w}, \phi, \lambda\}$, can be built up. The novel contribution of this work is to use this distribution of features to provide robust classification which properly accounts for the uncertainty in the estimated registration.

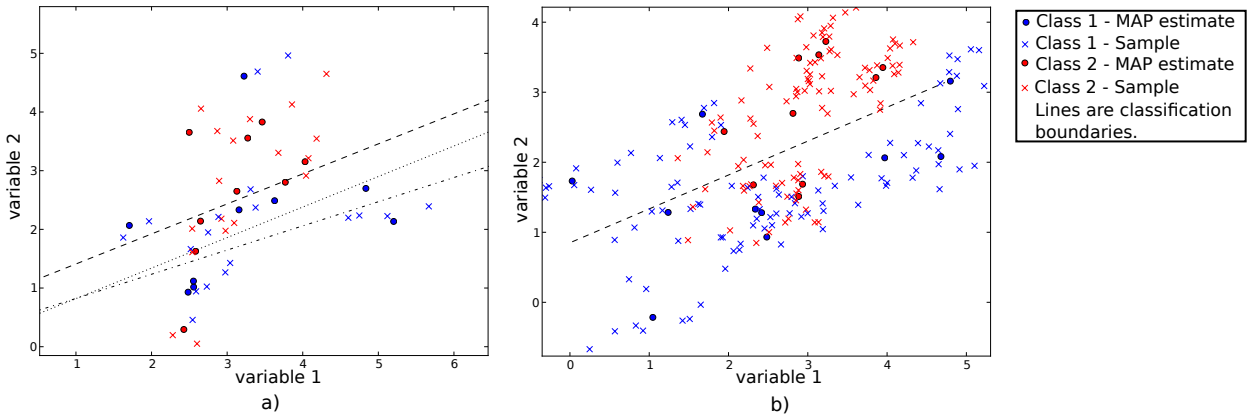


Figure 1: Graphical examples of using sampled data features in classification. a) illustrates ensemble learning with parametric bootstrapping (train+), each classification boundary is estimated using sampled data features from a fixed set of subjects. b) shows estimating variability in classification label using 10 random samples for each subject (test+).

Ensemble learning methods [2], are approaches which construct a set of multiple predictive models, and take an average of their predictions. In this case, we create several predictive models by drawing feature data samples for the set of training subjects in place of observations, in a parametric bootstrapping approach [3]. A graphical illustration of how multiple predictive models can be generated is illustrated in Figure 1 a). In this work we use an un-weighted averaging of predictions, but several more advanced approaches exist.

The predicted class variability can be estimated using a set of probable feature data for each subject, rather than the most likely observation. This is illustrated in Figure 1 b).

3 Experiments

162 subject images were taken from the ADNI database[5], 81 suffered from Alzheimer's disease (AD), the remaining 81 were age matched normal control subjects (NC). All the images were initially affinely aligned to the MNI 152 atlas space using 9 degrees of freedom, and resampled to 1mm isotropic resolution. To apply VBM or TBM, a suitable atlas image is required. An atlas image was created by iteratively registering 40 NC subjects to their average. This resulted in a sharp atlas image which is representative of the normal control population. Figure 2 illustrates how regions of interest were selected for left and right hippocampi using an enlarged bounding box.

The subject images were non-rigidly registered to the atlas image using a 5mm FFD knot spacing for TBM, and a 10mm spacing for VBM. Samples of probable mappings were drawn from Υ, these mappings were used to create samples of feature data. From these samples a subject feature mean, and covariance matrix was estimated. To allow storage of the covariance matrices, the feature data was sub-sampled by a factor of 4 to 1800 voxels.

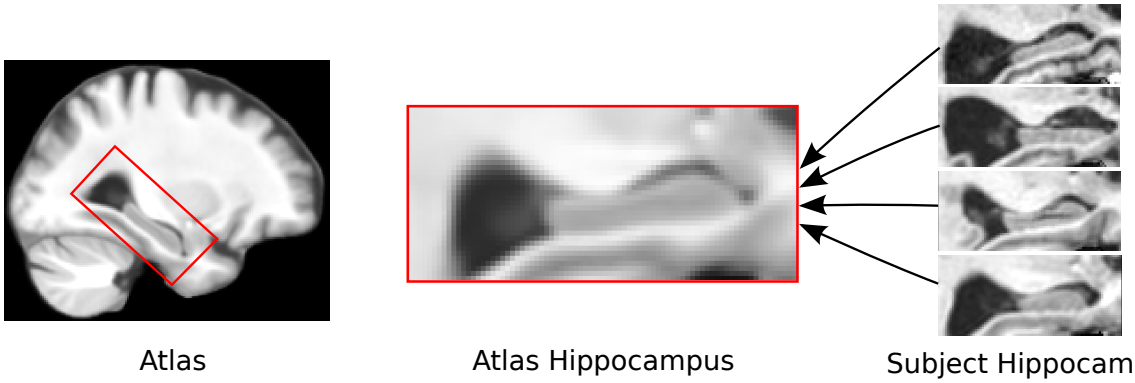


Figure 2: The experimental setup. A hippocampal bounding box of size (40x80x35 voxels), in red, selected the region of interest. Each subjects' hippocampi were registered to the atlas.

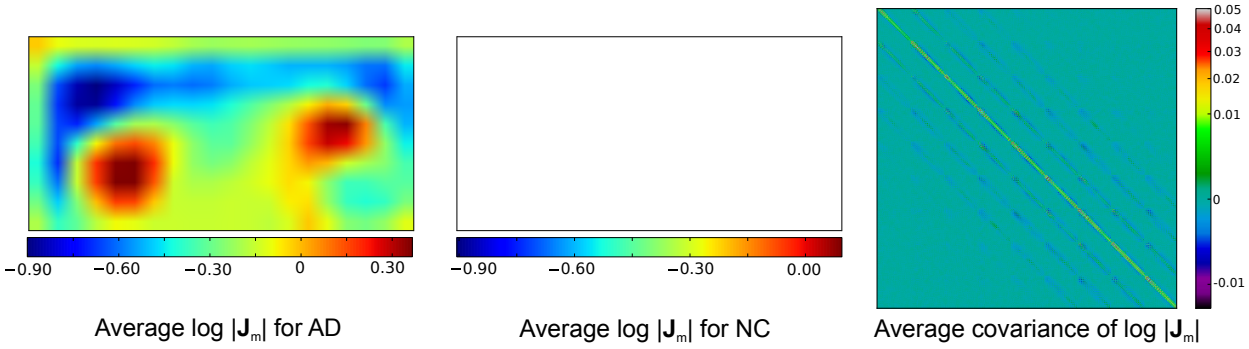


Figure 3: Left and middle images show the average $\log|\mathbf{J}_m|$ inferred from registering the subject to the atlas images, as shown in Figure 2, for the AD and NC groups. The image on the right shows the average of the intra-subject $\log|\mathbf{J}_m|$ covariance matrix as calculated based on the distribution of probable registration mappings.

The mean $\log|\mathbf{J}|$ images for both groups and the covariance matrix are given in figure 3.

Based on the parametric model of feature data for each subject, new probable samples can be drawn, and utilised as described in section 2.2. In the experiments we use a linear support vector machine (SVM), using all 1800 feature voxels to classify between subject groups. Subject age is regressed out for each voxel. We compare 4 method variants: original: where simply the subject mean image data is used; train+: where ensemble learning with parametric bootstrap is used to create multiple classifiers; test+: where extra test samples are classified; and finally: train+test+ which combines train+ and test+. Experimental results are given in table 1. For computational efficiency, all samples are drawn using the subject average covariance matrix. We also test averaging predictions across different feature types.

As shown in table 1, using train+, or train+test+ appears to provide a marked increase in

Table 1: Classification correct rate using the different predictor training and testing variants in TBM and VBM. L and R indicate the left and right hippocampus data, respectively.

Feature data	Original	Train+	Test+	Train+ Test+
L log $\|\mathbf{J}\|$	0.704	**0.796**	0.697	0.790
R log $\|\mathbf{J}\|$	0.709	0.759	0.722	**0.772**
Average L & R log $\|\mathbf{J}\|$	0.747	0.809	0.772	**0.821**
L VBM	0.759	**0.790**	0.759	**0.790**
R VBM	0.734	0.753	0.734	**0.759**
Average L & R VBM	0.778	**0.802**	0.772	0.784
Average L & R log $\|\mathbf{J}\|$ & VBM	0.802	**0.827**	0.802	0.809

classification rate over using the estimated MAP in both VBM and TBM approaches. The improvement in TBM is more substantial as TBM features are entirely dependent on the mapping used, whereas in VBM the feature data is less dependent on the mapping. Finally, we can see that all variants of ensemble learning provide reasonable estimates of classification uncertainty as the results from averaging multiple feature data types are improved.

4 Discussion and Conclusions

In this work we have demonstrated that using random sampled observations of spatially normalised feature data in statistical prediction leads to more robust prediction. The random feature data samples were derived from the set of probable mappings between subject and atlas space as inferred by a probabilistic registration tool. These samples were incorporated into an ensemble learning framework. In our experiments we provide results on the problem of classification of subjects with Alzheimers Disease, from age matched healthy controls using a linear SVM. The results show a consistent improvement in classification rate when using random samples compared to MAP observations, with a maximum improvement of 9% for some image derived features.

Possible future work includes using feature selection instead of voxel sub-sampling and investigating alternative ensemble learning approaches which use weighted averaging.

References

[1] J. Ashburner and K.J. Friston. Voxel-based morphometry–the methods. *Neuroimage*, 11 (6):805–821, 2000.

[2] T. Dietterich. Ensemble methods in machine learning. *Multiple classifier systems*, pages 1–15, 2000.

[3] B. Efron. Bootstrap methods: another look at the jackknife. *The annals of Statistics*, 7 (1):1–26, 1979.

[4] A. Klein and et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage*, 46(3):786–802, 2009.

[5] S.G. Mueller, M.W. Weiner, L.J. Thal, R.C. Petersen, C. Jack, W. Jagust, J.Q. Trojanowski, A.W. Toga, and L. Beckett. Alzheimer's Disease Neuroimaging Initiative. *Advances in Alzheimer's and Parkinson's Disease*, pages 183–189, 2008.

[6] P. Risholm, E. Samset, and W. Wells. Bayesian Estimation of Deformation and Elastic Parameters in Non-rigid Registration. *WBIR*, pages 104–115, 2010.

[7] D. Rueckert, LI Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using Free-Form Deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.

[8] I.J.A. Simpson, J.A. Schnabel, A.R. Groves, J.L.R. Andersson, and M.W. Woolrich. Probabilistic inference of regularisation in non-rigid registration. *NeuroImage*, 59(3): 2438–2451, 2012.

Image Restoration

Restoration of phase contrast microscopy images for the analysis of lung epithelial scratch wound repair assays

Joscelyn Sarsby[1]
jxs082@bham.ac.uk

Ela Claridge[2]
e.claridge@cs.bham.ac.uk

Gerard Nash[3]
g.nash@bham.ac.uk

David Thickett[3]
d.thickett@bham.ac.uk

Hannah Jeffery[3]
jefferHC@bham.ac.uk

ShengXing Zheng[3]
ZhengSY@bham.ac.uk

[1]PSIBS Doctoral Training Centre, School of Chemistry, University of Birmingham, Edgbaston, Birmingham B17 2TT.

[2]School of Computer Science, University of Birmingham, Edgbaston, Birmingham B17 2TT

[3]School of Clinical and Experimental Medicine, Collage of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham, B17 2TT

Abstract

In phase contrast microscopy living biological specimens can be studied without the need for staining or fluorescent tagging. This technique works by converting small differences in the optical path length of light passing through a transparent specimen into changes in image brightness. Objects appear typically dark surrounded by bright halos. The halo artefacts make automatic segmentation non-trivial. This paper develops a simplified model of image formation and its inversion implemented as a deconvolution with a Difference of Gaussians filter. In the restored images, values are proportional to the cell thickness and simple thresholding provides effective segmentation. This method is applied to segmentation and analysis of epithelial scratch wound repair assays used in basic bioscience research into treatment of acute lung injury. The computational method shows excellent agreement with the biosciences gold standard (correlation coefficient 0.92) whilst being objective and much faster.

1 Introduction

In phase contrast microscopy (PCM) small differences in the optical path length of light passing through a transparent specimen are converted into changes in image brightness. This makes it possible to image living biological specimens without the need for staining or fluorescent tagging. In PCM images background takes on mid-grey-level values (corresponding to little or no difference) whereas objects typically show as dark shapes surrounded by bright halos. The contrast between the dark interior and the bright halo is

related to *local* differences in specimen thickness. For this reason any global segmentation techniques such as thresholding cannot be straightforwardly applied.

Acute lung injury is a clinical condition that reduces the efficiency of the lungs to transfer gases in and out of the blood, causing hypoxia. There are no pharmacological approaches to promote the recovery and the current research aims to identify proteins and signalling pathways involved in the repair process to identify or develop pharmaceuticals that could help reduce the mortality rate. There are several models of repair of epithelial lung cells. The most common one is conceived as a two-phase process: of cell spreading and migration to cover the affected area, followed by cell proliferation to increase the cell density. An *in vitro* experimental procedure that is relatively straightforward to use is the wound repair assay. This method mimics the repair of cells. It comprises of growing a monolayer of cells that is then scratched to simulate wounding and loss of cells in the alveolus. Images are taken at predefined time points to calculate the percentage cell coverage of the wound, the rate of cell migration and the cell density [1]. As the epithelial cells are transparent, phase contrast microscopy is a method of choice.

Current gold standard protocol for processing the scratch wound images is to draw around the wound by hand. However, this method is time consuming, subjective, and can lead to bias [1]. Several computer image analysis methods have been applied to this problem. Holistic methods focus on textural or frequency differences between the scratch region and the regions populated by cells (e.g. [2]). Such methods work well for assessing cell migration but are not suitable for computing cell density. Methods, which are capable of this, rely on segmentation of individual cells. A simple approach is to apply thresholding followed by post-processing [3] which can often be quite considerable, for instance by evolving a geometric active contour around each cell candidate [4]. Watershed-based methods exploit the fact that the cell interior forms a “basin” whose edges define the boundaries of the cell. Over-segmentation is a common problem here although it can be overcome by careful pre-processing [5][6]. In a recently published method image restoration is carried out prior to segmentation. By modelling the optical pathway of a phase contrast microscope and its inversion the estimate of the optical path length is computed at every pixel [7]. In the resulting images the cells show as bright objects on a dark background and the halo effect is completely eliminated making segmentation easy.

This paper presents a method for assessing both the cell migration and the changes in cell density. Individual cells are segmented by applying a simplified inverse model of image formation in PCM. Individual segmented cells are then grouped on the basis of their proximity, subdividing the image into three regions: the central scratch area and the two confluent groups of cells on both sides of the scratch. The size of the scratch is computed for the central region and cell density is computed from the size of the areas occupied by cell and from the number of cells occupying it.

2 Methods

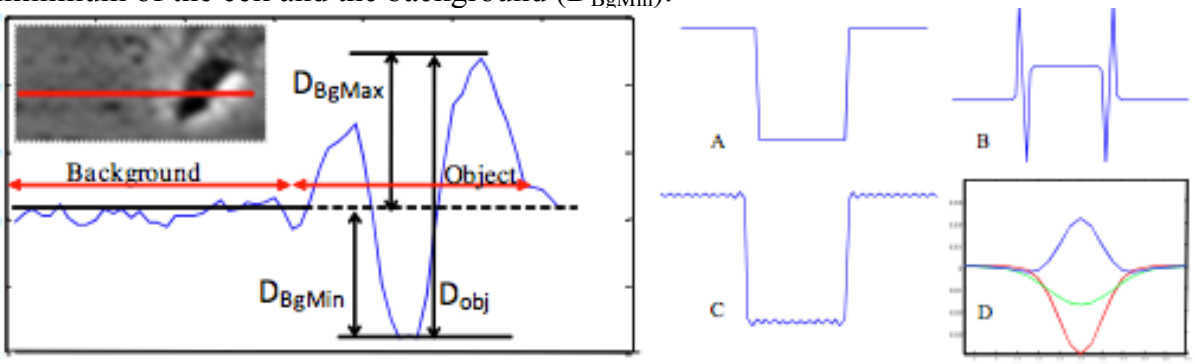
2.1 Simplified model of image formation

The two key processes responsible for the PCM image appearance are (1) the optical “computation” of local differences in optical path length and (2) the blur introduced by the microscope’s optics. In a gross simplification these processes combined can be represented by a difference of Gaussians (DoG). Figure 1 shows the profile through a PCM cell image and a profile of a wedge convolved with a DoG filter. A hypothesis is that by de-

convolving a PCM image with a DoG filter with suitably chosen parameters it is possible to compute relative estimates the optical path length for each point in the image. The parameters sought are a point spread function (PSF) for a given microscope and the magnitude of the difference between the two Gaussians, which is related to a phase shift.

2.2 Parameterisation of the Difference of Gaussians filter

This is a semi-manual process which needs to be performed only once for a given microscope and an imaging protocol. It requires the user to choose a profile though a single individual cell chosen to have a typical appearance. The profile must include a length of background. Three quantities are then computed from the profile (see Fig. 1): the difference between the maximum and minimum of the cell (D_{obj}), the difference between the maximum of the cell and the background (D_{BgMax}) and the difference between the minimum of the cell and the background (D_{BgMin}).



(a) (b)

Figure 1: (a) A single cell in a PCM image (inset) and its image profile; (b) (A) A wedge, (B) the wedge convolved with DoG filter, (C) restored edge: (B) deconvolved with the same DoG, (D) DoG filter (blue) generated by subtracting Gaussian (σ_1) (red) from Gaussian (σ_2) (green).

The parameters sought are the spread of Gaussian1, σ_1, the spread of Gaussian2, σ_2, and the difference $\Delta\sigma = \sigma_1 - \sigma_2$. To obtain values for σ_1 and σ_2 a 2D grid search was implemented. The axes of the grid are the two variables of the DoG, σ_1 and σ_2. The values on the grid are derived from eq. 1 when the sample image profile is deconvolved with the DoG parametrised by σ_1 and σ_2. The key problem was to formulate an objective function that uses only PCM image data. Initial research used synthetic objects where the DoG parameters were known and so the search could be directed (Fig. 2). By using parameters D_{obj}, D_{BgMax} and D_{BgMin} – all of which can be obtained from the PCM single cell image – it was observed that there was a distinct minimum in the value by using a specific combination (Eq. 1). The correct DoG occurs when the search is directed from different grid directions (vertical, horizontal and diagonal). The slope both before and after the minimum is the same so a gradient based method would be ineffective. Limits in the grid search were set for the ranges of both Gaussians and then every DoG was calculated and measured by the parameters in the equation:

$$Combination = D_{obj} \times \left(\frac{1}{D_{BgMin}}\right) \times L_{BgMax} \quad \text{(Equation 1)}$$

The DBgMax is not used directly within this calculation. Instead it is used as an indicator as to whether the DoG is too big or too small, which can cause the deconvolved image to be blurred; LBgMax is set to 0 or 1 to ensure that DBgMax does not become too large.
Large DoG also has smoothing effects and so D_{obj} is small which mimics an ideal situation.

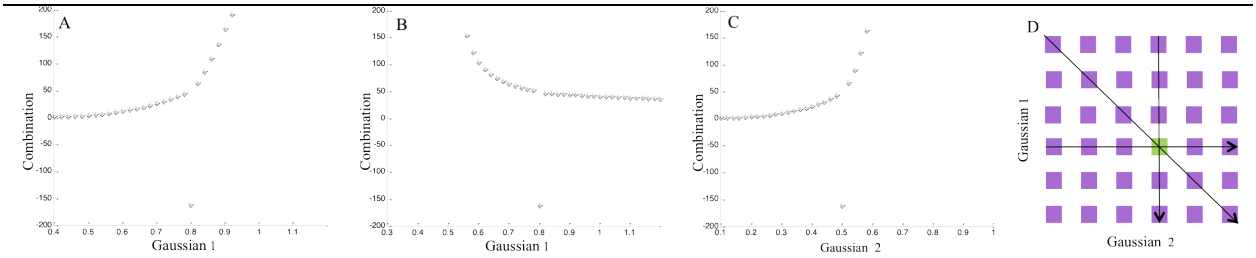


Figure 2: Directed search for the optimal difference of Gaussian. (A) Vertical search, (B) Diagonal search and (C) horizontal search. (D) shows how the directed search was made to overlap the correct combination of Gaussians (green square).

To evaluate the robustness of the recovery of parameters σ_1, σ_2, and $\Delta\sigma$ the process was repeated many times for different cells in the same image and for different images from the same batch of experiments. As these parameters characterise the imaging system they should show little difference across the experiments. The variation between the recovered parameters σ_1 and σ_2 was 0.073 and 0.054 respectively with a mean of 1.40 (σ_1) and 0.4 (σ_2) demonstrating the robustness of the process.

2.3 Deconvolution and post-processing

The phase contrast images were deconvolved with the DoG filter, $DoG=G(\sigma_1)-G(\sigma_2)$ yielding a relative estimate of cell optical pathlength which is proportional to cell thickness. The deconvolved images were observed to have large background variations (Fig. 3B). To remove them the original PCM image was strongly smoothed, deconvolved with the same DoG filter and subtracted from the restored original image (Figs. 3C and D). After the application of a top hat filter [8] (Fig. 3E) a single threshold value produced clear segmentation of the cells. To remove the artefacts the cell size and cell distance to neighbouring cells (obtained from distance transform) were computed and those segmented objects whose properties were outside one standard deviation from the respective mean were removed (Fig. 3F).

2.4 Derivation of quantitative parameters

By applying a sequence of morphological operations (dilation, erosion and hole filling with the structuring element of the size equal to half-diameter of a mean-size cell) to the binary segmented image three regions were obtained, corresponding respectively to the scratch wound and to the groups of cells on either side of the wound. The outline of the scratch wound is shown in red in (Fig. 3A). The manually drawn outline is shown in blue for comparison in Figure 3F. The cell count on each side of the wound divided by the segmented area provided an estimate of the cell density. The size of the central region provided an estimate of the scratch wound area.

3 Results and evaluation

Wound repair images are analysed in pairs, at zero and at an end time points (typically 18 or 24 hours). The ratios of the wound areas at the two time points are then computed. 22 pairs of images have been analysed both manually and by the method described. As the computer analysis was carried out retrospectively the actual outlines of the wound areas were not available in all the cases, only the wound area ratios for image pairs. For this data

there was good agreement between the manually and automatically derived ratios (correlation coefficient of 0.92, the slope of the trend line of 0.88). Reproducibility of the manually derived ratio is not known for the data sets used for these experiments. The latter could suggest that the method is more prone to error with pairs of images where the wound has closed to the greatest extent. In such samples at the end-time point the wound has partially closed and therefore the wound area is substantially more intricate. Overall trend between the two data sets shows that the gold standard produces a larger wound coverage value than the method described here. This could be due to errors in the boundary of the wound produced by this method or errors in bias in the gold standard.

The analysis has been carried out on two further experimental data sets of similar size and all the results showed similar correlation with the gold standard. In all the cases the same DoG filter was used, demonstrating the robustness of the imaging model and its derivation.

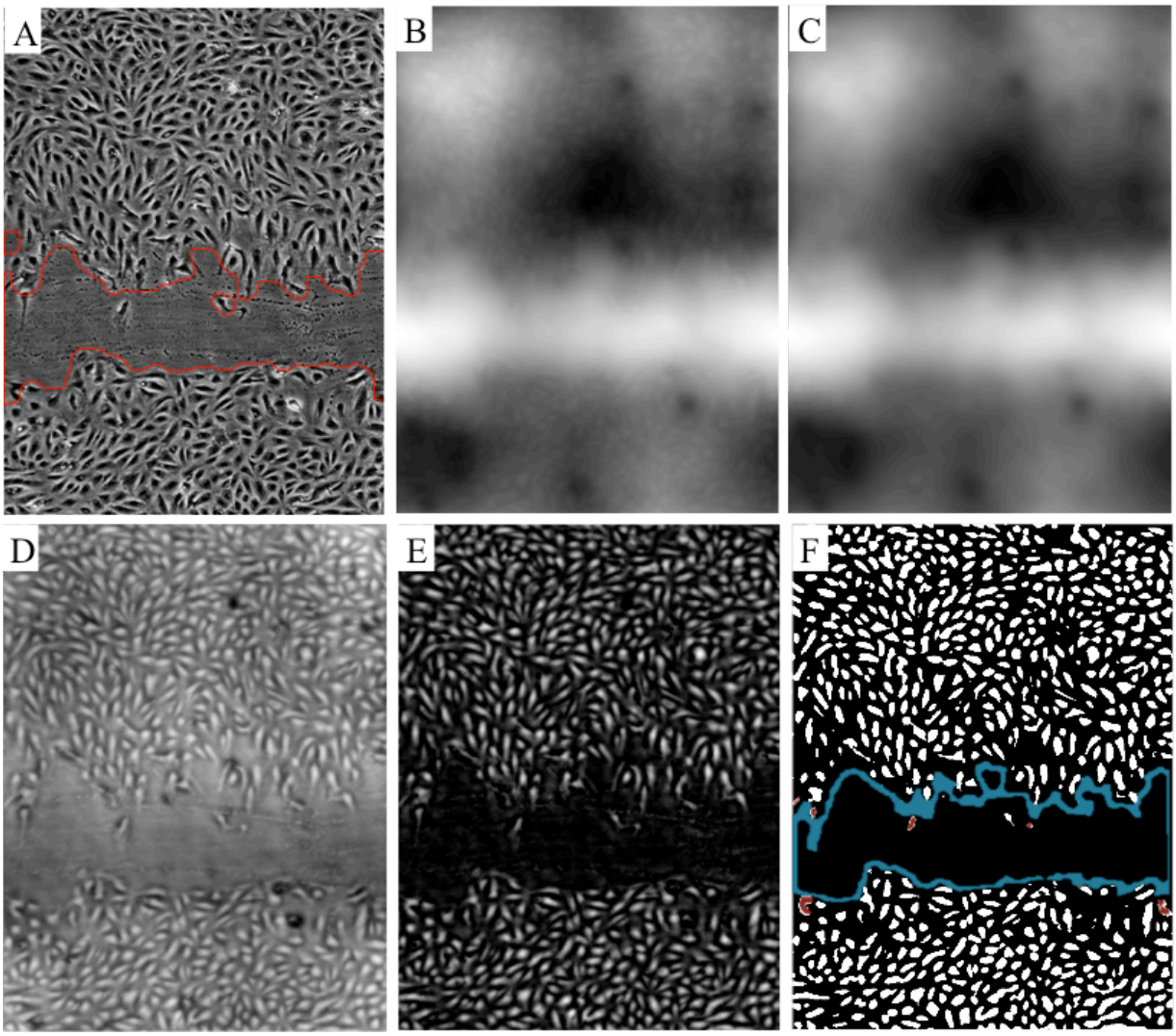


Figure 3: The Steps of the Method. (A) Initial phase contrast image with outline of automated segmentation produced by method described in this paper. (B) Image after deconvolution with DoG filter; (C) Background generated by deconvolving smoothed original image; (D) Image after background subtraction; (E) Top-hat algorithm applied to (D); (F) Binary image of (D) after thresholding; superimposed are cells that have been considered outliers and the outline of the wound when analysed by hand.

4 Discussion and Conclusion

The novel aspect of this work is the development, parameterisation and use of a model of image formation in the analysis of PCM images. The model itself is a very simple linear filter, DoG, lacking the nuance and quantitative aspects of more complex models (e.g. [7]). The key to its successful performance is its parameterisation, which has been computed via optimisation using a small image sample. Once the parameters have been established there is no regularisation involved in the process of image restoration as it is the case in [7]. The model inversion is implemented as a deconvolution. In the restored images the values are proportional to cell thicknesses, making cells easy to segment by simple global thresholding. Artefacts characteristic of the PCM, such as halos, are absent.

In application to the analysis of lung epithelial scratch wound repair assays this semi-automatic method has a number of advantages over the manual gold standard method. First is that by using this method the results are unbiased. Secondly, the method is significantly faster than the gold standard. Once a single cell profile has been determined (see 2.2) the method takes only 7 seconds to run on a standard PC using an uncompiled Matlab script. With the cell profile analysis time is increased to approximately 3 minutes, but this needs to be carried out only once per batch. In contrast, manual analysis takes approximately 10 minutes per image. The presented method can compute parameters related both to cell spreading and to cell proliferation. As the epithelial scratch wound repair assays are used in many areas of biosciences, and are applicable to the cell lines that can form monolayers [9], the authors plan to develop a practical tool for use by bioscientists.

References

[1] J.C. Yarrow et al., *BMC Biotechnology* 4 (21), 2004

[2] C.C. Reyes-Aldasoro et al., *Electronic Letters* 44 (13):791-793, 2008.

[3] C.J. Bradhurst et al., Segmentation of bone marrow stromal cells in phase contrast microscopy images. *23rd International Conference Image and Vision Computing*, New Zealand, IEEE, 1-6, 2008.

[4] M.E. Ambuhl et al., *Journal of Microscopy* 245 (2):161–170, 2012.

[5] O. Debeir et al Phase contrast image segmentation by weak watershed transform assembly., *International Symposium of Biomedical Imaging: From Nano to Marco*, France, 724-727, 2008

[6] J. Orikawa and T. Tanaka, Cell segmentation from phase-contrast images using hybrid watershed and region growing algorithm for genomic drug discovery. *Proceedings of SICE Annual Conference,* 84-88, 2010

[7] Z. Yin, et al., Understanding the Optics to Aid Microscopy Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2010*, 209-217. 2010

[8] L.G. Rodriguez et al., *Methods in Molecular Biology* 294, 23-29, 2005

[9] B. Coomber and A. Gotlieb, *Arteriosclerosis, Thrombosis, and Vascular Biology* 10 (2):215-222. 1990

Acknowledgments

The authors would like to acknowledge the EPSRC for funding.

Stack Alignment Transform for Misalignment Correction in Cardiac MR Cine Series

Constantine Zakkaroff[1]
mnkz@leeds.ac.uk
Aleksandra Radjenovic[2]
a.radjenovic@leeds.ac.uk
John Greenwood[2]
j.greenwood@leeds.ac.uk
Derek Magee[1]
d.r.magee@leeds.ac.uk

[1] School of Computing
University of Leeds
Leeds, LS2 9JT, UK

[2] School of Medicine
University of Leeds
Leeds, LS2 9JT, UK

Abstract

Stack misalignment in cardiac MR cine series distorts the correct appearance of anatomical features and reduces the reproducibility of volumetric measurements computed for diagnostic purposes. This paper describes a method for correction of stack misalignment in cardiac MR series. Our method involves registration against a reference volume and features a prominent enhancement designed to circumvent the weaknesses associated with slice-to-volume registration. The core of the presented method is a custom stack alignment transform which parametrises the in-plane movement for all slices independently of each other; at the same time the image similarity metric for every optimisation iteration is calculated on the whole stack with all slice correction parameters contributing to the result. The method was evaluated on 50 clinical datasets with a 100 simulated optimisation runs for every dataset. The results show that on average the method is able to recover alignment parameters with sub-voxel accuracy.

1 Introduction

Excellent tissue contrast along with high spatial and temporal resolution in Cardiovascular Magnetic resonance (CMR) imaging cast it as one of the most accurate and reproducible modalities for Coronary Heart Disease (CHD) diagnosis. Comprehensive CMR exams consist of a number of scans which provide in a single test the greatest volume of diagnostic data, including detailed ventricular and coronary anatomical information, myocardial function, perfusion, tissue viability and coronary blood flow. Wall motion and volumetric measurements for the assessment of myocardial function are derived from cine stacks spanning the cardiac cycle. Breath-holding is used as the means to exclude respiratory motion in cine series; a patient holds their breath multiple times and one to three slices over all phases are acquired during each breath-hold. Stack misalignment in cine series is caused by the inconsistencies in breath-hold positions between slice acquisitions. This paper presents a novel solution for cine correction which is motivated by the role of registration mediator that cine series play in spatiotemporal registration of MR perfusion and angiography.

1.1 Misalignment Correction Background

As with most registration applications, the solutions to the problem of slice misalignment in cine stacks rely on a reference image, which provides the bias for motion recovery during registration. A solution describing long-axis (LA) cine series as the reference image for short-axis (SA) slice-by-slice 2D translation misalignment correction in [7] provides reliable results; the use of 4D (3D + time) images in the registration enhances the appeal of the method, but it is only feasible if the LA series are multi-slice series; this is not the case in every study. Another solution relies on the slice-to-volume registration of SA slices to a high-resolution reference volume [2]; the multi-resolution registration optimises a 3D rigid transform for each SA slice from end-diastolic phase as it is registered to the whole heart volume. The results of stack correction with slice-to-volume registration experiments show the improvement in the accuracy for ventricular function analysis, although the study was performed on a relatively small number of healthy volunteers. Furthermore, as the authors in [1] observe, slice-to-volume registration can become an under-constrained problem due to the method's weakness which lies in out-of-plane rotations. A method with the potential to avoid the pitfalls of slice-to-volume registration in [6] was used for correction of multi-slice foetal brain MR images; the solution does not rely on slice-to-volume registration, however it requires at least three orthogonal stacks of slices.

Research publications on cardiac motion provide the evidence of through-plane respiration-induced motion of up to 23 mm translation along with some rotation (which varies considerably for each patient) due to patients' inconsistency in breath-hold positions [9]. Such displacements in the context of the method in [2] with its unconstrained 3D motion of the individual slices indicate that there is a high likelihood that some slices would be moved past their initial positions in the stack. While in the general context it is desirable for the individual slices to be registered to their precise 3D locations in the context of the reference volume, one side-effect of such correction is likely to manifest as some overlaps and "holes" in the reconstructed volumes, which in turn may require the use of some interpolation technique. However, the quality of reconstructions with linear interpolation based on non-adjacent stack slices degrades the features in the image.

1.2 Stack Alignment Transform Overview

Stack alignment transform has been developed as a component in an over-arching project aimed at establishing the co-location of perfusion defects (observable in perfusion series) with the coronary vessels (identifiable in angiography volumes) that might be responsible for these defects [11]. The project proposes a method of spatiotemporal registration of perfusion and angiography data; in this method the cine series are used for deriving the non-rigid transforms spanning an arbitrary difference between cardiac phases. The phase-to-phase transforms can be used to map the coronary anatomy derived from angiography volumes into an arbitrary phase of the cardiac cycle, *e.g.* the perfusion coordinate space. Hence the method for spatiotemporal registration is primarily concerned with recovering the finer components of the contractile motion such as the rotation of the myocardium and coronary vessels around the LA of the heart. In this context the primary concern is to constrain the slice displacements to in-plane motion to avoid the loss of image features due to out-of-plane slice displacements and non-adjacent slice interpolation. The parametrisation of the stack alignment transform deliberately avoids out-of-plane displacements; this approach presents a suitable alternative for cine correction in the context of perfusion and angiography fusion.

More generally, our solution is designed to circumvent the issues associated with slice-to-volume registration. Consider the case of slice-to-volume registration where the Mutual Information (MI) image similarity metric is evaluated from the region of overlap of a single slice and a volume: in this instance the joint probability distribution is likely to be a poor representation of the *true* joint probability distribution. However, a much better representation of the joint probability can be obtained from the overlap of the complete stack and reference volume because in this case the joint histograms computed for the calculation of the MI value contain a larger number of co-occurrences of grey-levels which represent the types of tissue in the images. This method of MI optimisation is made possible with a customised stack alignment transform presented and evaluated in this paper.

2 Materials and Methods

CMR images in this study were acquired on a dedicated cardiac research scanner (1.5 Tesla Intera CV, Philips, Best, The Netherlands) during a comprehensive trial on CMR imaging in CHD, CE-MARC [4]; the experiments involved two types of cardiac data:

Coronary MR angiography: Three dimensional, whole heart coronary MR angiography acquired using a balanced SSFP sequence and a respiratory navigator to compensate for respiratory motion during free breathing. Timing of the diastolic rest period is estimated from the four-chamber free breathing cine scan; matrix 304×304, field of view 320-460 mm, slice thickness 0.9 mm, 100-120 slices as required.

Resting wall-motion: Contiguous cine stack encompassing the entire left ventricle (LV) in 10-12 slices (depending on LV LA length) acquired during multiple breath-holds; matrix 192×192, field of view 320-460 mm, slice thickness 10 mm, at least 20 phases per cardiac cycle, $1-2$ slices per breath-hold.

Further information on the pulse sequences is available in [3]. For this study 50 representative datasets from the CE-MARC trial were selected randomly by the clinical trials unit; one dataset was excluded because of severe motion artefacts in the angiography volume.

2.1 Cine Phase Selection and Angiography Pre-registration

Prior to the optimisation of the stack alignment parameters it is necessary to determine the cardiac phase in the cine series to match the end-diastolic reference angiography volume. Although either visual examination of the cine series or ECG-based trigger delay normalisation can be used to calculate the matching phase, this study uses MI-based phase match determined by the best MI value computed for all of cine phases and angiography volume during pre-registration. In practice, the best match usually falls in a window of ±4 phases within the normalised trigger delay of the angiography volume. The pre-registration method in this case is a straightforward example of volume-to-volume registration with a translation transform, with a given cine phase and angiography volume used as fixed and moving images respectively. A set of elliptical masks derived from manually-defined rectangular regions of interest (one region for each slice) in the end-diastolic phase of the cine was used to exclude irrelevant anatomical features. Multi-resolution registration (two levels) with a Regular Step Gradient Descent optimiser and Matte's implementation of MI [8] available in ITK [5] provided fast and reliable results.

It should be noted that although it would be just as easy to optimise a rigid 3D transform with rotation instead of a translation-only transform, it was deliberately avoided. The

rationale for this decision was that the high-resolution angiography volumes carry the most amount of anatomical information in comparison to the sparse and likely misaligned cine series; theoretically there can be multiple cases of stack misalignment favouring 3D rotations. Hence it is preferable to use the simplest possible transform to avoid rotational errors.

2.2 Stack Alignment Transform and Misalignment Correction

Our stack alignment transform is a custom spatial transform which parametrises the in-plane movement along the X and Y dimensions for the individual stack slices independent of each other. In addition, the transform also includes a parameter for global translation along the Z direction. Fixed transform parameters include image origin, spacing and size, all obtained from the corresponding cine phase. For example, if the cine series consist of 12 slices, the transform will be parametrised as $T = \{S_1, \ldots, S_{12}, z_g\}$, where $S_n = \langle x_n, y_n \rangle$ and n is the slice number calculated from the fixed parameters based on a given point in 3D space. For a point $P = [x, y, z]$ in 3D space, the transformed point $P' = [x', y', z']$ is calculated as follows:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x + x_n \\ y + y_n \\ z + z_g \end{bmatrix} \tag{1}$$

Apart from the transform, the rest of the registration components for this stage are the same as for pre-registration: multi-resolution pyramids (two levels), Regular Step Gradient Descent optimiser, Matte's MI and fixed image masks. During the registration the cine stack is used as the fixed image, because the image MI similarity metric collects samples from the moving image on the fixed image grid. After the registration the inverted transform can be applied to correct all phases in the cine series.

2.3 Evaluation Protocol

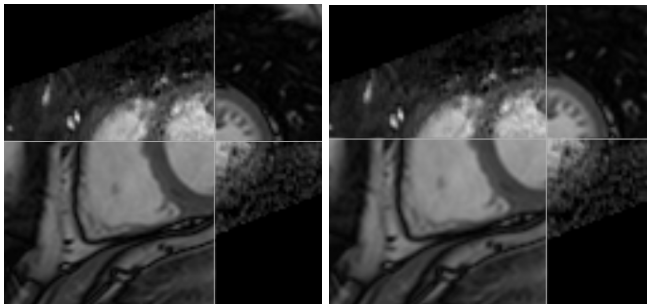
The performance of alignment parameter recovery was evaluated with a protocol of registration uncertainty measurement described in [10]; in this evaluation method the mean transform is used as a pseudo-gold standard based on the assumption of a zero mean distribution of errors. The core of the protocol is in the assessement of the accuracy of registration in a large number of runs each with a known introduced misalignment. In our study a 100 random simulated misalignments for all slices in a given dataset were generated; the in-plane offsets $X_M = \{x_{M1}, \ldots, x_{M100}\}$ and $Y_M = \{y_{M1}, \ldots, y_{M100}\}$ were restricted to ± 4 mm. This type of simulated error is significantly more severe than the misalignment common for clinical datasets, because typically only a small portion of slices in a stack contains observable movement, while the simulated transforms displace every slice. During the evaluation, a clinical dataset with its unknown error repeatedly was further misaligned with one of the 100 simulated offsets. If after registration with a 100 simulated offsets the recovered in-plane displacement for each slice are to be represented as $X_R = \{x_{R1}, \ldots, x_{R100}\}$ and $Y_R = \{y_{R1}, \ldots, y_{R100}\}$ respectively, then the estimated correction parameters for the slice are represented as $X = X_M - X_R$ and $Y = Y_M - Y_R$ with their arithmetic means $\bar{X}$ and $\bar{Y}$ representing the mean transform over 100 simulated runs. The magnitude of registration error distance $D_{err}(n)$ for slice n shifted with a simulated offset n can be calculated as follows:

$$D_{err}(n) = \sqrt{(x_n - \bar{X})^2 + (y_n - \bar{Y})^2} \tag{2}$$

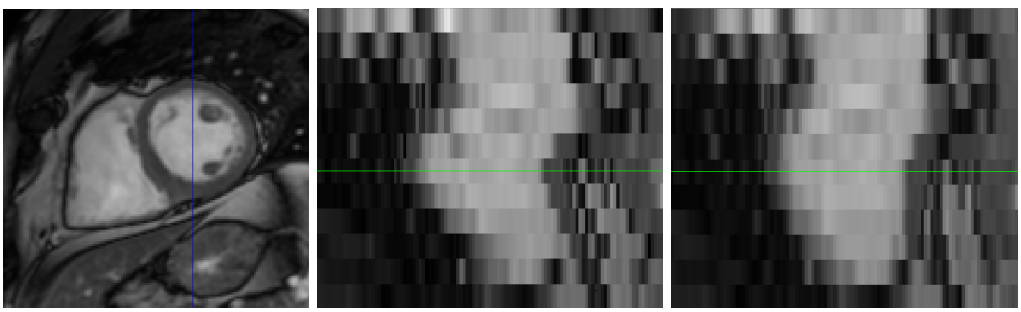
3 Results

Visual inspection of the registration results for the 49 datasets (without simulated errors) showed that alignment was recovered successfully in all cases; Checkerdoards showing cine and double-oblique reformatted reference volumes were used for visual validation. Figure 1 presents a pair of examples of misalignment and its correction (typical for this study); most datasets needed correction for a small subset of slices (usually one to three) while a few datasets required correction for all slices. Although no obvious registration failures were detected it should be noted that correction of apical slices may not always be viable due to the absence of clearly visible myocardium and blood-pool. A small number of apparent phase mismatch cases possibly could be explained by the unaccounted out-of-plane rotations.

Parameter recovery using the stack alignment transform with simulated offsets described in the previous section was compared against the slice-by-slice registration (restricted to in-plane movement) with the same parameters. Figure 2 shows the comparison of the distributions of the error distance (as defined in equation (2)) for the 49 datasets; the higher error ranges for the extreme apical and basal slices are explained by three factors: (a) insufficient image features (*e.g.* no clearly visible ventricular/myocardial walls); (b) insufficient overlap region between apical or basal slices in cine and angiography; (c) presence of a rotational component of displacement which is not considered in this study. For the slices with clearly visible myocardium, the error mean is around one voxel (1.3 mm approx.) and is bounded by the 2.5 mm in 95% of cases for both slice-by-slice alignment and stack alignment experiments. However the difference between the overall shape of the error distributions shows that stack alignment transform is more consistent in all slices and has a greater capture range.



(a) SA before (b) SA after



(c) Slicing plane indication (d) LV before (e) LV after

Figure 1: Misalignment correction examples: checkerboards in (a) and (b) show the alignment of a single SA slice from a cine stack and double-oblique reformatted angiography volume before and after correction; images in (d) and (e) show one phase from a cine stack before and after correction; orientation and location of the slicing plane is indicated in (c).

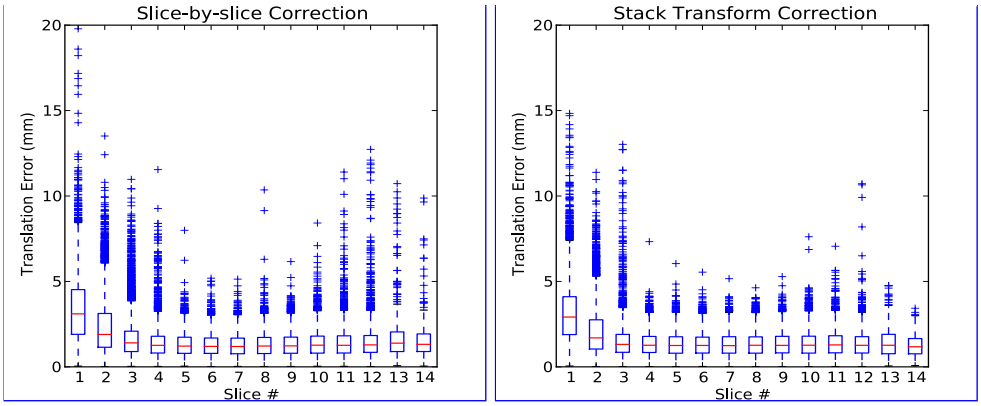


Figure 2: Distribution of error distance D_{err} as defined in equation (2); left: slice-by-slice correction; right: stack alignment transform correction. The boxplots summarise the errors for a 100 simulated runs for all datasets, one bar per single slice index in a stack.

4 Discussion and Conclusions

In this paper we presented an alternative solution for stack misalignment correction in MR cine series. The solution relies on a novel spatial transform, the stack alignment transform; it has been developed in the context of spatiotemporal registration of perfusion and angiography data, where it is preferable to avoid the degradation of the image features in the stack in order to recover the finer components of the contractile motion reflecting the radial rotation around the LA of the heart. The key feature of the stack alignment transform is its ability to harness the power of the MI image similarity metric by enabling the computation of the joint probability distributions based on the overlap of the whole stack and reference volume, rather than an overlap of one slice with the reference volume. The evaluation of the stack alignment transform on a large clinical dataset proves that the proposed method for cine correction offers a robust solution which restricts the single slice motion to in-plane translation and favours the integrity of the cine series. In future work we anticipate extending the stack alignment transform to recover the rotational component of the myocardial motion. In addition, a more extensive evaluation with manually drawn myocardial contours to be used as the gold-standard could provide a further insight into the benefits of the stack alignment transform; in our experience, visual correction validation with checkerboards is not specific enough for assessing the finer components of recovered motion.

Acknowledgements

Clinical data were obtained within the CE-MARC study funded by the British Heart Foundation Grant to Leeds CMR Unit (grant RG/05/004–JPG, AR). This work was funded by the Top Achiever Doctoral Scholarship awarded by Tertiary Education Commission of New Zealand (grant UOLX08001–CZ) and WELMEC, a Centre of Excellence in Medical Engineering funded by the Wellcome Trust and EPSRC (grant WT 088908/Z/09/Z–AR, DRM).

References

[1] W Birkfellner, M Figl, J Kettenbach, J Hummel, P Homolka, R Schernthaner, T Nau, and H Bergmann. Rigid 2D/3D Slice-to-volume Registration and its Application on Fluoroscopic CT Images. *Medical Physics*, 34(1):246–255, 2007.

[2] A Chandler, R Pinder, T Netsch, J A Schnabel, D J Hawkes, D L G Hill, and R Razavi. Correction of Misaligned Slices in Multi-slice Cardiovascular Magnetic Resonance Using Slice-to-volume Registration. *Magnetic Resonance Imaging*, 10(1):13, 2008.

[3] J P Greenwood, N Maredia, A Radjenovic, J M Brown, J Nixon, A J Farrin, C Dickinson, J F Younger, J P Ridgway, M Sculpher, S G Ball, and S Plein. Clinical Evaluation of Magnetic Resonance Imaging in Coronary Heart Disease: The CE-MARC Study. *Trials*, 10(1):62–71, 2009.

[4] John P Greenwood, Neil Maredia, John F Younger, Julia M Brown, Jane Nixon, Colin C Everett, Petra Bijsterveld, John P Ridgway, Aleksandra Radjenovic, Catherine J Dickinson, Stephen G Ball, and Sven Plein. Cardiovascular Magnetic Resonance and Single-photon Emission Computed Tomography for Diagnosis of Coronary Heart Disease (CE-MARC): A Prospective Trial. *Lancet*, 379(9814):453–60, 2012.

[5] L Ibáñez, W Schroeder, L Ng, and J Cates. *The ITK Software Guide*, 2nd edition, 2005.

[6] Kio Kim, Piotr a Habas, Francois Rousseau, Orit a Glenn, Anthony J Barkovich, and Colin Studholme. Intersection Based Motion Correction of Multislice MRI for 3-D *in Utero* Fetal Brain Image Formation. *IEEE Transactions on Medical Imaging*, 29(1): 146–158, 2010.

[7] J Lötjönen, M Pollari, S Kivistö, and K Lauerma. Correction of Movement Artefacts from 4-D Cardiac Short- and Long-axis MR Data. In *Medical Image Computing and Computer Assisted Intervention*, 2004.

[8] D Mattes, D R Haynor, H Vesselle, T K Lewellyn, and W Eubank. Nonrigid Multimodality Image Registration. In *Medical Imaging: Image Processing*, volume 4322 of *SPIE Conference Series*, pages 1609–1620, 2001.

[9] K McLeish, D L G Hill, D Atkinson, J M Blackall, and R Razavi. A Study of the Motion and Deformation of the Heart Due to Respiration. *IEEE Transactions on Medical Imaging*, 21(9):1142–50, 2002.

[10] J R Sykes, D S Brettle, D R Magee, and D I Thwaites. Investigation of Uncertainties in Image Registration of Cone Beam CT to CT on an Image-guided Radiotherapy System. *Physics in Medicine and Biology*, 54(24):7263–83, 2009.

[11] C Zakkaroff, D Magee, A Radjenovic, and R Boyle. Mediated Spatiotemporal Fusion of Multiple Cardiac Magnetic Resonance Datasets for Patient-specific Perfusion Analysis. In *Computing in Cardiology*, volume 37, pages 469–472, 2010.

Deformable Models & Energy Minimisation

Fractional Entropy Based Active Contour Segmentation of Cell Nuclei in Actin-Tagged Confocal Microscopy Images

Leila Meziou[1]
leila.meziou@ensea.fr
Aymeric Histace[1]
aymeric.histace@u-cergy.fr
Frédéric Precioso[2]
precioso@polytech.unice.fr
Bogdan J. Matuszewski[3]
BMatuszewski1@uclan.ac.uk
Franck Carreiras[4]
franck.carreiras@u-cergy.fr

[1] ETIS - UMR 8051 - CNRS
University of Cergy-Pontoise/ENSEA,
Cergy-Pontoise, France
[2] I3S - UMR 6070 - CNRS
University of Nice/Sophia-Antipolis,
Nice, France
[3] ADSIP Research Centre
University of Central Lancashire
Preston, UK
[4] ERRMECe
University of Cergy-Pontoise
Cergy, France

Abstract

In the framework of cell structure characterization for predictive oncology, we propose in this paper an unsupervised statistical region based active contour approach integrating an original fractional entropy measure for single channel actin tagged fluorescence confocal microscopy image segmentation. Following description of statistical based active contour segmentation and the mathematical definition of the proposed fractional entropy descriptor, we demonstrate comparative segmentation results between the proposed approach and standard Shannon's entropy obtained for nuclei segmentation. We show that the unsupervised proposed statistical based approach integrating the fractional entropy measure leads to very satisfactory segmentation of the cell nuclei from which shape characterization can be subsequently used for the therapy progress assessment.

1 Introduction

Segmentation of cellular structures is an essential tool in cell microscopic imaging as it enables measurements which can be used to track cell divisions or help to reconstruct corresponding cell lineage trees providing data for calculation of different parameters like cell proliferation rate for instance. More specifically, the work presented in this paper has been carried out in a context of analyzing changes of cell cytoskeleton properties in a response to ionizing radiation insult. The final goal of this research effort is to better understand cell bio-mechanical responses during cancer radiation therapy. In this context, we propose an unsupervised segmentation approach of fluorescence confocal microscopy images which

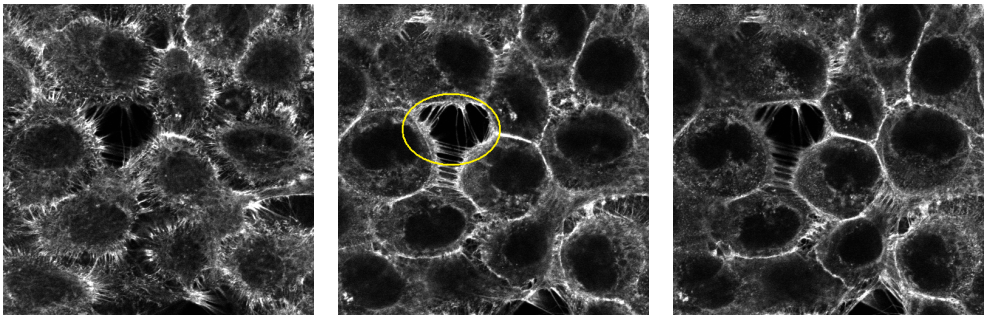
represents practical computational problem when considering many monolayer acquisitions – in order to effectively extract nuclei as a first step for providing spatial reference frame for analyzing cytoskeleton changes.

To date, only few methods have been proposed to address direct segmentation (without any denoising preprocessing of acquired images) of cell structures in fluorescence confocal microscopy images. In former approaches proposed in [8] and [12], authors focused on nuclei segmentations. In [14], authors proposed cell segmentation in 2D-fluorescent images with two channels (actin and nucleus tagging) using a multiphase level-set combining Chan-Vese [2] and geodesic active contour models, together with repulsive force introduced to prevent segmented cells from overlapping. In [7, 15] automated 3D cell segmentation from a 3D confocal acquisition of early Zebrafish embriogenesis is proposed; Two different fluorescent markers (red for nuclei and green for membrane) are used to easily discriminate nuclei from cell membranes. In [15], authors introduced an adapted version of the subjective surface technique [13] for surface reconstruction from missing boundary information whereas [7] use a multiphase level-set based on probability correlation functions.

Within a level set framework as in [7, 14], our method aims at a different objective: segmentation of microscopic 2D images extracted from a full single channel confocal acquisition with only one fluorescent marker used for actin tagging. The filament actin (F-actin) is believed to play a vital role in cell structure [3]. As Actin is one of the three most common proteins in human cytoskeleton, analysing its changes and properties could be instrumental in analysis of cell properties. For example this can be associated with cancer evolution. Nevertheless, due to a highly complex actin appearance, a high level of noise and a strong non-homogeneity of intensity and gradient information, the segmentation of cell structures in such imaging data, is a very challenging task. Moreover, a particular attention is given to completely avoid any enhancement preprocessing [10] and to reduce to its minimum, manual interventions during the whole segmentation process.

The data used in this paper were obtained from human prostate cells (PNT2). Actin were labelled with phalloidin-FITC and all imaging was carried out using a Zeiss LSM510 confocal microscope. Fig. 1 shows different slices from the microconfocal acquisition of the monolayer PNT2 cell culture. The stack volume is defined on the $512 \times 512 \times 98$ grid of pixels each $0.21\ \mu\text{m} \times 0.21\ \mu\text{m} \times 0.11\ \mu$ in size. Actin is mostly present at the periphery of the cells and within the cytoplasm. We can then notice that high intensities of actin tagged confocal images allows biologists to roughly delineate cell membranes whereas darkest areas are identified as nuclei. Owing to the high level of Poisson noise corrupting these images and the particular texture of actin, classic region based active contour approach, like the Chan and Vese one [2], fails even in segmenting properly the boundaries of nuclei corresponding to each cell [6]: We then propose to tackle this segmentation using statistical based active contour (see [5] for an overview on the work on this area) more adapted to this particular context than classic region based ones.

The remaining of this article is organized as follows: in Section 2, the framework of statistical based active contour using entropy estimation is first presented, subsequently the corresponding Partial Differential Equation (PDE) which steers the evolution of the contour is described and finally attention is focused on the particular proposed fractional entropy descriptor; Section 3 focuses on experiments on microscopic images followed by conclusions and perspectives drawn in Section 4.



(a) (b) (c)

Figure 1: Examples of actin tagged fluorescence confocal microscopy images extracted from a 3D microconfocal acquisition of the monolayer PNT2 cell culture. (a) Lower slice (with low z-stack index), (b) Mid-slice with the lowest level of structural noise (a "hole" is highlight in yellow which should not to be confused with a nucleus), (c) Upper slice with non-homogeneity of the fluorescent marker on the left hand side.

2 Active contour segmentation using a fractional entropy descriptor

Formerly introduced in [1], integral based features active contour methods are derived from traditional region based approaches by utilizing integral statistics as descriptors of the inner (Ω_{in}) and outer (Ω_{out}) regions delimited by the active curve Γ at a given iteration τ of the segmentation process.

First, let $H(\Omega_i)$ denotes an integral entropy estimation associated to a particular region Ω_i within image such as

$$H(\Omega_i) = \int_{\Omega_i} \varphi\left(\hat{p}(\mathbf{I}(\mathbf{x}), \Omega_i)\right) d\mathbf{x}\,, \tag{1}$$

with φ a monotonic increasing function, $\mathbf{I}(\mathbf{x})$ the luminance of pixel $\mathbf{x} = (x, y)$ and $\hat{p}$ the non-parametrically estimated Probability Density Function (PDF) of region Ω_i. More precisely, usually PDF $\hat{p}$ is estimated using Parzen window technique such as:

$$\hat{p}(\mathbf{I}(\mathbf{x}), \Omega_i) = \frac{1}{|\Omega_i|} \int_{\Omega_i} G_\sigma(\mathbf{I}(\mathbf{x}) - \lambda)\, d\mathbf{x}, \tag{2}$$

where $\lambda \in [0 \ldots 2^n - 1]$, n is the quantization level of image intensity function, and G_σ is the Gaussian kernel of standard deviation σ. In the framework of statistical region based active contour segmentation, corresponding functional to minimize H_T is defined as a competition between inner and outer regions characterized by the considered entropy descriptor H of Eq. (1) such as:

$$H_T = H(\Omega_{in}) + H(\Omega_{out}) + g \int_\Gamma ds, \tag{3}$$

where g is a positive real value and s standard arclength of the curve. This functional combines measures of the considered entropy descriptor of inner Ω_{in} and outer Ω_{out} regions of the curve for a given iteration τ of the segmentation with an additional regularization term minimizing the curve length. The Euler derivative of Eq. (3) and usual minimization scheme leads to the Partial Differential Equation (PDE) steering the evolution in the orthogonal direction $\mathbf{N}$ of the active curve Γ :

$$\frac{\partial \Gamma}{\partial \tau} = \big(A(s, \Omega_{in}) + \varphi(p(\mathbf{I}(s), \Omega_{in})) + A(s, \Omega_{out}) + \varphi(p(\mathbf{I}(s), \Omega_{out})) + g\big)\mathbf{N} \tag{4}$$

where A is related to the proposed descriptor and is defined by:

$$A(s,\Omega_i) = -\frac{1}{|\Omega_i|}\int_{\Omega_i} \varphi'(\hat{p}(\mathbf{I}(\mathbf{x}),\Omega_i))[\hat{p}(\mathbf{I}(\mathbf{x}),\Omega_i) - G_\sigma(\mathbf{I}(\mathbf{x}) - \mathbf{I}(s))]\,d\mathbf{x}\,. \tag{5}$$

For illustration, let's consider the particular case of Shannon's entropy: φ function is given by $\varphi(r) = -r.log(r)$ and then $H(\Omega_i) = -\int_{\Omega_i} \hat{p}(\mathbf{I}(\mathbf{x}),\Omega_i) log(\hat{p}(\mathbf{I}(\mathbf{x}),\Omega_i))d\mathbf{x}$. As it will be shown in the Experiment section, standard Shannon's entropy have some limitations in terms of segmentation performance: more specifically, this measure makes segmentation of corrupted (Gaussian noise) textured images challenging [4], in the case of high level of structural noise, the segmentation results are not that satisfactory. This can be explained by the fact that Shannon's entropy tassumes that the corrupting noise (and then the corresponding PDF $\hat{p}$) can be parametrically modeled within the exponential family [5] which is not true when considering confocal microscopy data. In this particular context, fractional entropy like the Rényi, given by:

$$H_R(\Omega_i) = \frac{1}{1-\alpha} log \int_{\Omega_i} \hat{p}(\mathbf{I}(\mathbf{x}),\Omega_i)^\alpha\, d\mathbf{x}\,, \tag{6}$$

is of primary interest. Rényi's entropy shows some relaxation possibilities regarding the shape of the PDF $\hat{p}$ which can be used by a judicious setting of α (considered as strictly positive and lower than one in this study [11]). Unfortunately, Rényi's entropy as expressed in Eq. (6) is part of the non-integral entropy family that can not be easily associated to a region-based criterion in a classic active contour based segmentation. Nevertheless, taking benefits of the properties of Rényi's entropy, we propose to define a fractional entropy measure adapted to the framework of statistical region-based active contour segmentation. For this, let consider Eq. (1) with φ function and its derivative given by:

$$\varphi(r) = -\frac{1}{1-\alpha} log\,(r^\alpha) \;\text{ and }\; \varphi'(r) = -\frac{\alpha}{(1-\alpha)r}\,. \tag{7}$$

Considering φ function of Eq. (7), we obtain an integral entropy measure integrating a fractional parameter.

3 Experiments and results on nuclei segmentation

In this section, comparative segmentation results obtained are first described for the unsupervised nuclei segmentation within the mid-slice of the considered single channel confocal microscopic acquisition (Fig. 1(b)).

The PDE from Eq. (4) is implemented in the level-set framework in order to be able to automatically handle topological changes [9]. The initialization of the active contour is a set of small circles uniformly distributed all over image which allows an easy initialization of the algorithm. Classic AOS scheme is used for implementation in order to obtain a reasonably fast convergence segmentation.

Fig. 2 shows results obtained with the standard Shannon's entropy criterion and the proposed fractional entropy descriptor. Considering experiments based on Shannon's entropy (Fig. 2 (left)), as one can notice, the method does not lead to satisfying results . Figs. 2 (middle and right) shows results of nuclei segmentation on the same slice, but with the proposed fractional entropy criterion: the nuclei segmentation is definitely improved. As one can notice, as actin is a complex structure, some artifacts could appear. It is possible to

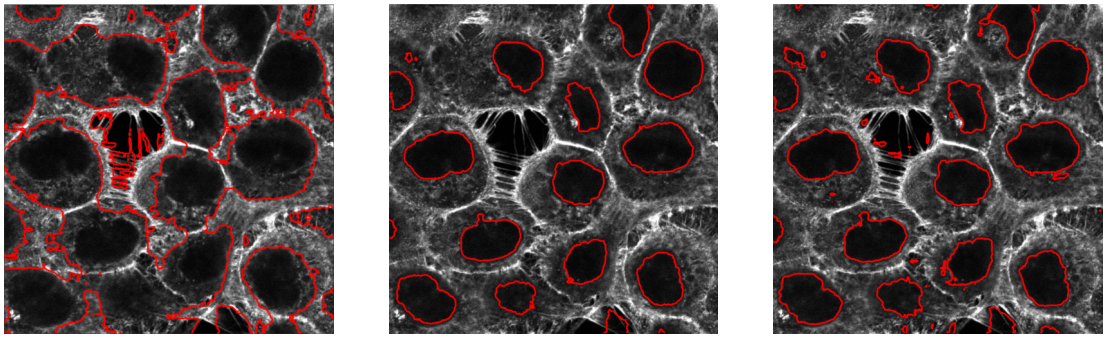


Figure 2: Comparative results of nuclei segmentation. Left: Shannon's entropy ; Middle and Right: Fractional entropy descriptor with $\alpha = 0.5$ (middle) and $\alpha = 0.7$ (right) ; for all experiments $g = 10$.

overcome this drawback with an adapted choice of α parameter. As one can see in Fig. 2, for $\alpha = 0.5$, smaller number of artifacts related to α value and those results show that this parameter plays an important role in the sensitivity of the criterion to the level of corrupting noise. Moreover, it is important to notice that the proposed fractional entropy measure can also distinguish a hole from a nucleus (which method based on Shannon's criterion was not able to achieve), whereas the associate PDFs are statistically very similar.

Finally, Fig. 3 shows some segmentation results obtained on the whole stack of acquired images. Results shown are obtained with $\alpha = 0.5$, and $g = 10$. To obtain these results, a propagation initialization strategy, starting on middle slice is used which makes integration of some spatial coherence within the segmentation scheme to avoid propagation of false detection due to complex appearance of actin.

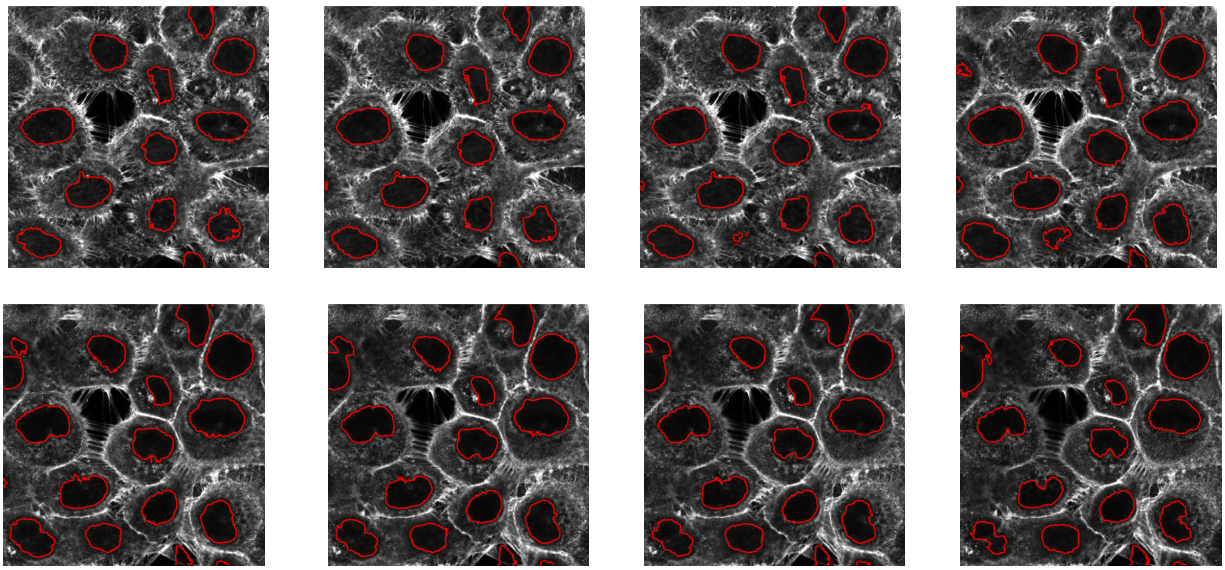


Figure 3: Segmentation of nuclei made on upper (upper row) and lower (bottom row) slices of the all stack, mid-slice of Fig. 2 being the initialization level). $\alpha = 0.5$ and $g = 10$.

These results have been qualitatively considered as very satisfactory from an expert point of view and a very good start for further investigations on that particular data.

4 Conclusion

The contribution of the segmentation approach presented in this article is twofold: (i) An unsupervised cell nuclei segmentation method is proposed for single channel actin tagged acquisitions without any enhancement or denoising preprocessing of the considered images. (ii) Whereas in the framework of statistical based active contour methods standard Shannon's entropy is most often considered as the region descriptor, we proposed an original fractional entropy measure inspired from Rényi's entropy making possible a relaxation of the sensibility of the descriptors to strong variations of the shapes of the non parametrically estimated related PDF. Main motivation was to overcome the limitations of Shannon's entropy which appeared not adapted to our segmentation problem. We are currently working on locally relating the optimal choice of α parameter with the level of noise and/or the type of texture characterizing the image to segment. From an application point of view, we are finalizing a 3D version of our slice-by-slice segmentation approach to have a direct visualization of the 3D shape of the nuclei. Membrane segmentations will be the next step.

Acknoledgements

This work was supported by the UK Engineering and Physical Sciences Research Council [TeRaFS project, grant number EP/H024913/1].

References

[1] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson. Image segmentation using active contours: Calculus of variations or shape gradients? *SIAM Journal on Applied Mathematics*, 63:2128–2154, June 2003.

[2] T. F. Chan and L. A. Vese. Active contours without edges. *IEEE transactions on Image Processing*, 10(2):266–277, February 2001.

[3] A. Hall. *The cytoskeleton and cancer*, volume 28. Springer Netherlands, Philadelphia, PA, USA, 06 2009. ISBN 0167-7659.

[4] A. Herbulot, S. Jehan-Besson, S. Duffner, M. Barlaud, and G. Aubert. Segmentation of vectorial image features using shape gradients and information measures. *Journal of Mathematical Imaging and Vision*, 25(3):365–386, 2006.

[5] F. Lecellier, M.J. Fadili, S. Jehan-Besson, G. Aubert, M. Revenu, and E. Saloux. Region-based active contours with exponential family observations. *Journal of Mathematical Imaging and Vision*, 36(1):28–45, January 2010.

[6] L. Meziou, A. Histace, F. Precioso, B. Matuszewski, and M. Murphy. Confocal Microscopy Segmentation Using Active Contour Based on Alpha-Divergence. In *Proceedings of IEEE ICIP conference*, pages 3138–3141, 09 2011.

[7] K. Mosaliganti, A. Gelas, A. Gouaillard, R. Noche, N. Obholzer, and S. Megason. Detection of spatially correlated objects in 3d images using appearance models and coupled active contours. In *Proceedings of MICCAI*, pages 641–648, London, UK, 2009. Springer-Verlag. ISBN 978-3-642-04270-6. doi: http://dx.

doi.org/10.1007/978-3-642-04271-3_78. URL http://dx.doi.org/10.1007/978-3-642-04271-3_78.

[8] C. Ortiz De Solorzano, E. Garcia Rodriguez, A. Jones, D. Pinkel, J. W. Gray, D. Sudar, and S. J. Lockett. Segmentation of confocal microscope images of cell nuclei in thick tissue sections. *Journal of Microscopy*, 193(3):212–226, 1999.

[9] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Comp. Phy.*, 79:12–49, 1988.

[10] S. Pop, A. Dufour, and J-C. Olivo-Marin. Image filtering using anisotropic structure tensor for cell membrane enhancement in 3d microscopy. In *Proceedings of IEEE-ICIP*, pages 2085–2088, September 2011.

[11] A. Rényi. On measures of entropy and information. *4th Berkeley Symposium on Mathematical Statistics and Probability.*, 1:547–561, 1960.

[12] A. Sarti, C. Ortiz De Solorzano, S. Lockett, and R. Malladi. A geometric model for 3d confocal microscope image analysis. *IEEE Transactions on Biomedical Engineering*, 47:1600–1609, 2000.

[13] A. Sarti, R. Malladi, and J. A. Sethian. Subjective surfaces: A geometric model for boundary completion. *International Journal on Computer Vision*, 46(3):201–221, February 2002. ISSN 0920-5691.

[14] P. Yan, X. Zhou, M. Shah, and S. T. C. Wong. Automatic segmentation of high throughput rnai fluorescent cellular images. *IEEE Transactions on Information Technology in Biomedicine*, 12(1):109–117, January 2008.

[15] C. Zanella, M. Campana, B. Rizzi, C. Melani, G. Sanguinetti, P. Bourgine, K. Mikula, N. Peyriéras, and A. Sarti. Cells segmentation from 3d confocal images of early zebrafish embryogenesis. *IEEE transactions on Image Processing*, 19(3):770–781, March 2010. ISSN 1057-7149.

Medical Image Segmentation Using Magnetostatic Active Contours (MAC) with Tensor Diffusion

Huaizhong Zhang
http://www.cs.swan.ac.uk/~hzhang/
Xianghua Xie
http://www.cs.swan.ac.uk/~csjason/

Department of Computer Science
Swansea University
UK, SA2 8PP

Abstract

In medical imagery, traditional deformable models often face substantial challenges due to fine structures and image complexity. Recently, based on magnetostatic theory, a new deformable model, namely MAC, is proposed for improving the ability of the active contour in dealing with complex geometries and segmentation difficulties. A Laplacian diffusion scheme is proposed in the MAC model to tackle excessive image noise which can interrupt image gradient vectors and in turn affect the external force field. In this paper, a derived vector potential field (VPF) is employed to obtain magnetic force and thus a diffusion tensor can be applied to diffuse VPF in terms of both magnitude and directional information, instead of directly diffusing the magnetic field as in the MAC model. Our diffusion is carried out both in spatial and temporal aspects of VPF so that the performance of the deformable model is significantly improved while images are with low signal-noise ratio (SNR) and poor contrast. In addition, the proposed diffusion enhancement can lead to evolving the curve smoothly and thus level set evolution is adapted to approach genuine object of interest. By applying in several medical image modalities, the results demonstrate the effectiveness of the proposed method.

1 Introduction

Due to natural biological variability and pathological conditions in medical imaging, various challenging problems arise in image segmentation, particularly low signal-noise ratio (SNR) and poor contrast. Over the decades, a substantial amount of methods have been developed to deal with these segmentation problems. Among many others, methods such as conventional active contour models, e.g. snake [6], gradient vector flow snake (GVF) [11] and geodesic active contours (GAC) [2], Markov random field (MRF) models [4], graph cuts [1], and piecewise constant models, e.g. Chan-Vese model [3], have been widely applied to medical image segmentation. It is generally perceived that region based approaches are more robust towards image noise and artefacts compared to edge based ones. However, region based approaches may suffer from inhomogeneity in regional characteristics, e.g. intensity or texture. Methods that have the properties of both approaches may offer a better solution in certain applications, particularly where object boundaries can not be simply described as discontinuities in intensity or regional characteristics.

Active contour model is a promising technique to medical image segmentation, due to its ability to handle complex object geometries. Numerous works have been reported in the literature. Explicit active contour models are usually represented by using parameterized splines such as snake [6], GVF [11], which are inadaptable to topological changes; in contrast, implicit active contour models are to embed the contour into a higher dimensional function so that these models can temporally adapt to contour propagation and topological changes. The level-set representation of the deforming contour has proved a powerful technique for numerically implementing the implicit active contour models such as GAC [2], the Chan-Vese model [3]. However, in medical imagery, these methods often face substantial challenges due to fine structures and image complexity, particularly convergence issues such as deep concavities, weak edges and broken boundaries. Recently, physics-inspired deformable models have been proposed for coping with these difficulties. For example, a charged particle model (CPM) [5] based on electrostatics was applied to localize object boundaries by assigning opposite charges to edge pixels and free particles. Another example is the elastic interaction based snake [9] which a long range interaction force based on the elastic interaction between line defects in solids is used as an external force in active contours. However, these methods still meet difficulties in dealing with the mentioned problems because weak edges can result in broken contours and noise can lead to curve evolution in wrong directions. Most recently, a new formulation for active contours based on magnetostatic field, called MAC, was introduced by Xie and Mirmehdi [10]. Instead of assigning fixed charges, MAC allows the charges flow through the edges and then a magnetic field is generated by the charge flow. Thus, the active contour is attracted towards the edges under the magnetic influence. Although the model is derived from simplistic edge based assumption, i.e. object boundary collocates with intensity discontinuity, the magnetic field computed from the interactions of those image gradient vectors behaves very similarly to a region based force. The MAC model shows significant improvements on the previous active contour models.

However, inevitably the MAC model will suffer from extensive image noise interference, which can disturb the gradient vectors and cause the inaccurate computation of the magnetic flux that will in turn affect curve evolution and lead to the deviation of the contour from genuine object boundaries. In [10], the authors proposed a Laplacian diffusion scheme to refine the magnetic field before contour evolution. In this paper, a derived vector potential field (VPF) is employed to obtain magnetic force and a diffusion tensor is then applied to diffuse VPF in terms of both magnitude and directional information. Thus, the diffusion is carried out both in spatial and temporal aspects of VPF so that the performance of MAC can significantly be improved for images with poor contrast and low SNR. In addition, the proposed diffusion enhancement can lead to evolving the curve smoothly and thus level set evolution is adapted to approach genuine object of interest. We apply the proposed method in two medical image modalities, angiography in eyes and urinary cast. The experimental results demonstrate the effectiveness of the proposed method. The remainder of this paper is organised as follows: In Section 2, a derived VPF is introduced and the calculation of magnetic flux with VPF is presented, and then tensor diffusion scheme for VPF is described. In Section 3, we present the experimental results in various medical image datasets. Finally, a conclusion is given in Section 4.

2 Method

2.1 Vector potential field

As described in [10], the direction of the currents, flows of charges, running through object boundary is estimated based on edge orientation, which is obtained by a 90° rotation in the image plane of the normalized image gradient vectors $(\hat{I}_x, \hat{I}_y)$, where I denotes an image. The image plane is considered as the X-Y plane in a 3D space Ω whose origin coincides with the origin of the image coordinates. Thus, the direction of object boundary current, $\mathbf{O}(\mathbf{x})$, is estimated as:

$$\mathbf{O}(\mathbf{x}) = (-1)^{\lambda}(-\hat{I}_y(\mathbf{x}), \hat{I}_x(\mathbf{x}), 0), \tag{1}$$

where $\mathbf{x}$ denotes a pixel position in the image domain, $\lambda = 1$ gives an anti-clockwise rotation in the image coordinates, and $\lambda = 2$ provides a clockwise rotation. In terms of the level set representation, the direction of current for the active contour, denoted as Υ, is similarly obtained by rotating the gradient vector $\nabla\Phi$ of the level set function Φ. Let $f(\mathbf{x})$ be the magnitude of image gradient, the magnetic flux $\mathbf{B}(\mathbf{x})$ generated by gradient vectors at each $\mathbf{x}$ is computed as:

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \sum_{\mathbf{s}\neq\mathbf{x}} f(\mathbf{s})\mathbf{O}(\mathbf{s}) \times \frac{\hat{\mathbf{R}}_{\mathbf{xs}}}{R^2_{\mathbf{xs}}}, \tag{2}$$

where μ_0 is the permeability constant, $\mathbf{s}$ denotes an edge pixel position, $\hat{\mathbf{R}}_{\mathbf{xs}}$ denotes a 3D unit vector from $\mathbf{x}$ to $\mathbf{s}$ in the image plane, and $R_{\mathbf{xs}}$ is the distance between them. The active contour is assigned with unit magnitude of electric current. The force imposed on it is derived as:

$$\mathbf{F}_m(\mathbf{x}) \propto \Upsilon(\mathbf{x}) \times \mathbf{B}(\mathbf{x}). \tag{3}$$

The magnetostatic active contour (MAC) model is then formulated as:

$$C_t = \alpha g(\mathbf{x})\kappa\hat{\mathbf{N}} + (1-\alpha)(\mathbf{F}_m(\mathbf{x}) \cdot \hat{\mathbf{N}})\hat{\mathbf{N}}, \tag{4}$$

where $g = 1/(1+f)$, κ denotes the curvature, and $\hat{\mathbf{N}}$ is inward unit normal. Note, $\mathbf{F}_m$ lies in the image domain and its third element equals zero, which can be ignored.

Furthermore, the magnetic flux density $\mathbf{B}$ shown in (2) can be described by its magnetic vector potential $\mathbf{A}(\mathbf{x})$:

$$\mathbf{B}(\mathbf{x}) = \nabla \times \mathbf{A}(\mathbf{x}), \mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \sum_{\mathbf{s}\neq\mathbf{x}} f(\mathbf{s}) \frac{\mathbf{O}(\mathbf{s})}{R_{\mathbf{xs}}}. \tag{5}$$

where $\mathbf{A}(\mathbf{x})$ can be expressed as $(A_i(\mathbf{x}), A_j(\mathbf{x}), 0)$ in Ω:

$$A_i(\mathbf{x}) = \frac{\mu_0}{4\pi} \sum_{\mathbf{s}\neq\mathbf{x}} f(\mathbf{s}) \frac{-\hat{I}_y(\mathbf{s})}{R_{\mathbf{xs}}}, \; A_j(\mathbf{x}) = \frac{\mu_0}{4\pi} \sum_{\mathbf{s}\neq\mathbf{x}} f(\mathbf{s}) \frac{\hat{I}_x(\mathbf{s})}{R_{\mathbf{xs}}}, \tag{6}$$

where we consider $\lambda = 1$ (see (1)). It does not make any theoretical difference if $\lambda = 2$ is used, which simply leads to $\mathbf{B}(\mathbf{x}, \lambda = 2) = -\mathbf{B}(\mathbf{x}, \lambda = 1)$. $(A_i, A_j, 0)$ is referred to as *vector potential field* or VPF for convenience. Equation (5) indicates that we can indirectly refine the magnetic field $\mathbf{B}$ by diffusing VPF, instead of post-processing $\mathbf{B}$ as proposed in [10] which, as we show in the experimental section, is problematic in noisy situations.

2.2 Tensor diffusion for vector potential field

For the vector potential field, referring to nonlinear diffusion techniques introduced in [8], we propose a tensor diffusion of the potential field before computing its circulation density. Thus, we diffuse the VPF by considering not only their magnitude but also their orientations

and then the magnetic flux $\mathbf{B}$ is calculated using the refined VPF according to (5). This will efficiently alleviate the disturbance of image noise or artefacts because the Laplacian diffusion used in [10] merely performs the scalar process in the magnetic flux $\mathbf{B}$. The following tensor diffusion scheme is employed for our purpose:

$$\frac{\partial}{\partial t}u - \nabla \cdot (\mathbf{D}(\nabla u)\nabla u) = \mathscr{F}(u_0), \tag{7}$$

where $u(t,\mathbf{x})$ is the diffused version, t can be considered as the "scale parameter", $\mathbf{D} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is the diffusion tensor (a positive definite symmetric matrix), $\mathscr{F}$ can be considered as a penalty function which forces the diffusion result to conform to certain criteria, and $u_0(\mathbf{x}) = u(0,\mathbf{x})$ denotes the initial state. In our case, the VPF $\mathbf{A} = (A_i, A_j, 0)$ is a vector field. Thus, the nonlinear diffusion takes the following coupled form:

$$\begin{cases} \frac{\partial}{\partial t}\mathscr{A}_i - \nabla \cdot (\mathbf{D}(\nabla \mathscr{A}_i, \nabla \mathscr{A}_j)\nabla \mathscr{A}_i) = \mathscr{F}(A_i), \\ \frac{\partial}{\partial t}\mathscr{A}_j - \nabla \cdot (\mathbf{D}(\nabla \mathscr{A}_i, \nabla \mathscr{A}_j)\nabla \mathscr{A}_j) = \mathscr{F}(A_j), \end{cases} \tag{8}$$

where $\mathscr{A}_i(0,\mathbf{x}) = A_i(\mathbf{x})$, $\mathscr{A}_j(0,\mathbf{x}) = A_j(\mathbf{x})$. The diffusion tensor can be decomposed into two orthogonal components, one of which is parallel to the local potential vector and the other is perpendicular to the local vector. The orientation of a vector in the potential field can be denoted as $(\cos\theta, \sin\theta)$ and its orthogonal unit vector can be obtained as $(-\sin\theta, \cos\theta)$. Thus, (8) can be re-written as:

$$\frac{\partial}{\partial t}\mathscr{A} - \nabla \cdot \left(\mathbf{R}^T \begin{pmatrix} \omega & 0 \\ 0 & \gamma \end{pmatrix} \mathbf{R}\nabla \mathscr{A} \right) = \mathscr{F}(\mathbf{A}), \tag{9}$$

where $\mathbf{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, ω is the diffusion function in the direction of the VPF and γ denotes the diffusion function orthogonal to the field. Note the divergence and gradient operations are applied to each spatial component of $\mathscr{A}$ separately.

Considering the fact that we aim to have larger diffusion where potential vectors have smaller magnitude and preserve large potential vectors that are spatially consistent, we select the weighting function ω: $\omega(\mathbf{A}(\mathbf{x})) = e^{-\frac{|\mathbf{A}(\mathbf{x})|^3}{K'}}$, where K' is the parameter controlling the amount of diffusion ($K' = 0.25$ in our experiments) and $|\mathbf{A}| \in [0,1]$. In addition, the diffusion perpendicular to local potential vectors plays a critical role in propagating the potential vectors from strong edges to regions further away from them, which may be dominated by image noise. We wish to increase the diffusion in this perpendicular direction according to the magnitude of the VPF, thus we choose the square root of the VPF magnitude, $|\mathbf{A}|^{0.5}$, as a measure of degradation along this perpendicular direction and then similarly define the diffusion function $\gamma = e^{|\mathbf{A}|^{0.5}}$.

For obtaining the solution with a nontrivial steady state in (9) and avoiding the problem of choosing a stopping time [7], we define the conformity function $\mathscr{F}$ as $\mathscr{F}(\mathbf{A}(\mathbf{x})) = \mathbf{A}(\mathbf{x}) - \mathscr{A}(\mathbf{x})$. Moreover, a weight is used to exert substantial diffusion for desirable regions, i.e. the larger ω is, the less constraint is imposed on conformity. Thus, $\mathscr{F}$ is given as:

$$\mathscr{F}(\mathbf{A}(\mathbf{x})) = (1 - \omega(\mathbf{A}(\mathbf{x})))(\mathbf{A}(\mathbf{x}) - \mathscr{A}(\mathbf{x})). \tag{10}$$

3 Experimental results

The proposed diffusion method has been applied in refining the magnetic flux of the MAC model for extracting the ROIs in medical images. The controlling parameter K' is set to

0.25 in the experiments. We compare this indirect diffusion method with the Laplacian diffusion method used in [10]. The first image modality used is from the angiography in eyes, which is usually with low SNR and poor contrast. The initialisations of two cases are presented in Fig. 1(a)(d). The results in Fig. 1(c)(f) illustrate that our method can extract more accurate boundaries of the vessels in comparison to the Laplacian method (Fig. 1(b)(e), see the arrowed places). Fig. 2 illustrates the curve evolution of an example using our method and the Laplacian. Three more examples in Fig. 3 show the superior performance of the proposed diffusion method in comparison with the Laplacian method. We can see that the acquired boundaries using the proposed method are not only more accurate, but also more smoother than using the Laplacian method. This is due to performing the tensor diffusion in terms of both the magnitude and orientation information of VPF.

Furthermore, we have applied our method to another medical image modality, urinary cast, which are often with low SNR and poor contrast as well. Four examples are presented in Fig. 4 where the largest components are the interested cellular structures in the images. The initialisations are shown in Fig. 4(a1)(b1)(c1)(d1). Fig. 4(a2)(b2)(c2)(d2) show the results using the Laplacian diffusion and Fig. 4(a3)(b3)(c3)(d3) present the results using our method. In this difficult situation, the proposed diffusion method can acquire accurate smooth boundaries while the Laplacian diffusion works by contrast ineffective.

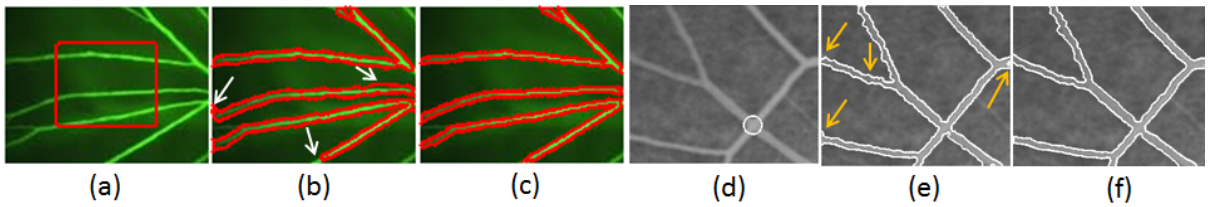


Figure 1: Segmentation results for two angiography images. (a,d) are the initialisations, (b,e) are the results using the Laplacian diffusion. (c,f) are the results using the proposed method. The arrowed places are problematic.

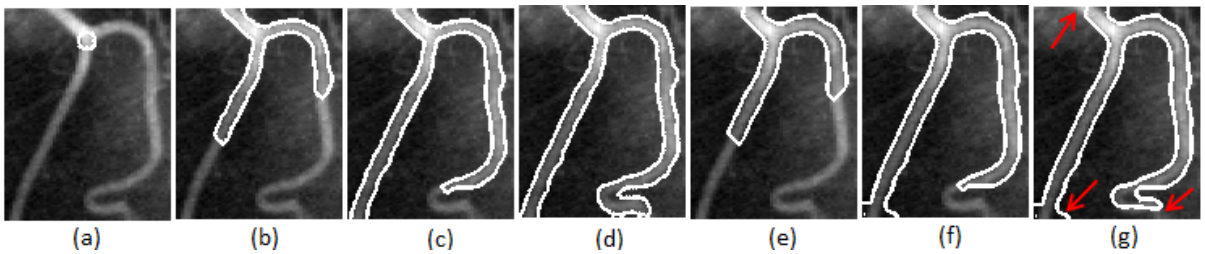


Figure 2: Curve evolution using our method. (a) is the initialisation; (b,e), (c,f) and (d,g) are the intermediate results for iteration 40, iteration 90 and the final one using our method and the Laplacian respectively. The arrowed places are problematic

4 Conclusions

In this paper, we propose a tensor diffusion method to diffuse the derived VPF so that the magnetic flux **B** can be refined better than the Laplacian method used in [10]. The major flavor of the proposed diffusion method is due to considering both diffusion magnitude and orientations in the diffusion process that can significantly improve the segmentation performance of the MAC model, particular in images with low SNR and poor contrast. We apply the proposed method in several medical image modalities. The current experimental results illustrate its superior performance in comparison to the Laplacian diffusion method. Future work will focus on the performance evaluation of the proposed method.

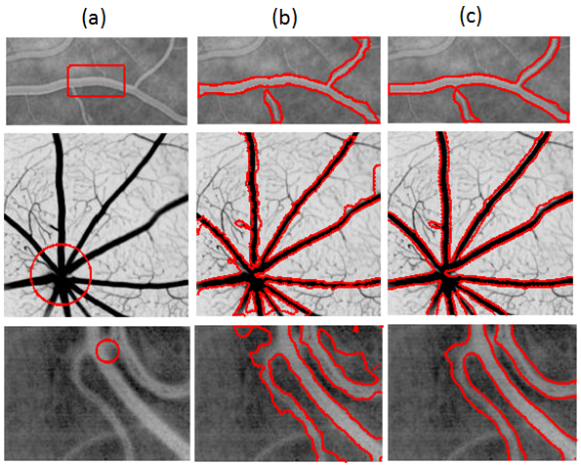


Figure 3: Extra examples for angiography. Column (a) is the initialisation, Column (b) is the results using the Laplacian method and Column (c) is the results using the proposed method.

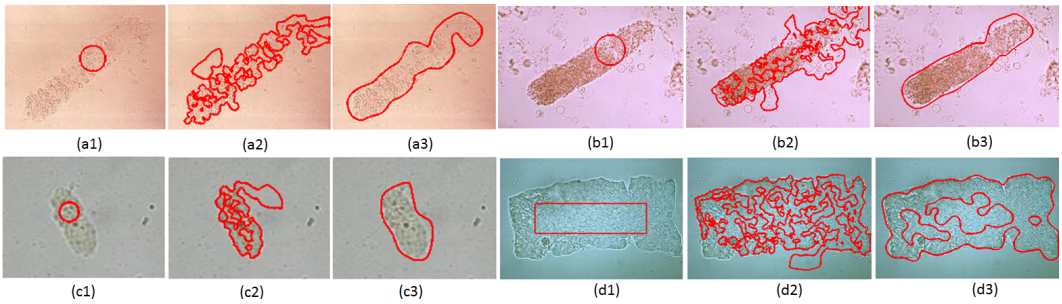


Figure 4: Results for urinary cast. (a1,b1,c1,d1) are the initialisations, (a2,b2,c2,d2) are the results using Laplacian diffusion, (a3,b3,c3,d3) are the results using the proposed method.

References

[1] Y. Boykov and G. Funka-Lea. Graph cuts and efficient n-d image segmentation. *IJCV*, 70(2): 109–131, 2006.

[2] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contour. *IJCV*, 22(1):61–79, 1997.

[3] T. Chan and L. Vese. Active contours without edges. *T-IP*, 10(2):266–277, 2001.

[4] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE. Tran. Pattern Anal. Mach. Intell.*., 6(6):721–741, 1984.

[5] A. Jalba, M. Wilkinson, and J. Roerdink. CPM: A deformable model for shape recovery and segmentation based on charged particles. *T-PAMI*, 26:1320–1335, 2004.

[6] M. Kass, A. Witkin, and D. Terzopoulus. Snakes: Active contour model. *IJCV*, 1(4):321–331, 1988.

[7] K. Nordström. Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection. In *ECCV*, pages 18–27, 1990.

[8] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner-Verlag, Stuttgart, Germany, 1998.

[9] Y. Xiang, A. Chung, and J. Ye. An active contour model for image segmentation based on elastic interaction. *J. Comp. Phy.*, 219(1):455–476, 2006.

[10] X. Xie and M. Mirmehdi. MAC: Magnetostatic active contour. *T-PAMI*, 30(4):632–646, 2008.

[11] C. Xu and J. Prince. Snakes, shapes, & gradient vector flow. *IP*, 7(3):359–369, 1998.

Unsupervised Tumour Segmentation in PET Based on Active Surface Modelling and Alpha Matting

Ziming Zeng[1,2]
zzz09@aber.ac.uk
Reyer Zwiggelaar[1]
rrz@aber.ac.uk

[1]Department of Computer Science
Aberystwyth University, Aberystwyth, UK
[2]Faculty of Information and Control Engineering
Shenyang Jianzhu University, Shenyang, China

Abstract

This paper presents a novel, unsupervised method for segmenting tumours in PET data. The method uses region-based active surface modelling in a hierarchical scheme to eliminate segmentation errors, followed by an alpha matting step to further refine the segmentation. We have validated our method on real PET images of head-and-neck cancer patients as well as custom designed phantom PET images. Experiments show that our method can generate more accurate segmentation than some previous approaches.

1 Introduction

Positron emission tomography (PET) is widely used for tumour imaging in cancer diagnosis, staging, treatment evaluation and radiotherapy planning, which require accurate segmentation of the volume of interest (VOI). Segmentation of PET VOIs is commonly performed by a combination of manual delineation and intensity thresholding. Manual segmentation is time-consuming and highly variable. Intensity thresholding lies at the core of most fully- or semi-automatic VOI segmentation methods [9]. More advanced methods include the Poisson Gradient Vector Flow (PGVF) of Hsu et al. [2] and the Markov Random Field Expectation Maximisation (MAP-MRF EM) labelling technique of Gribben et al. [4]. In this paper we propose a new VOI segmentation method for PET, which employs a hierarchical approach combining an improved region-based active surface model and alpha matting.

2 Image Segmentation Method

Our hierarchical segmentation scheme consists of three steps. (see Fig. 1). In the first two steps, an improved 3D active surface modelling method is used to segment the VOIs. In the last step, a trimap which contains a definite foreground, a definite background and an unknow region is automatically generated and an alpha matting technique is used to refine the segmentation results.

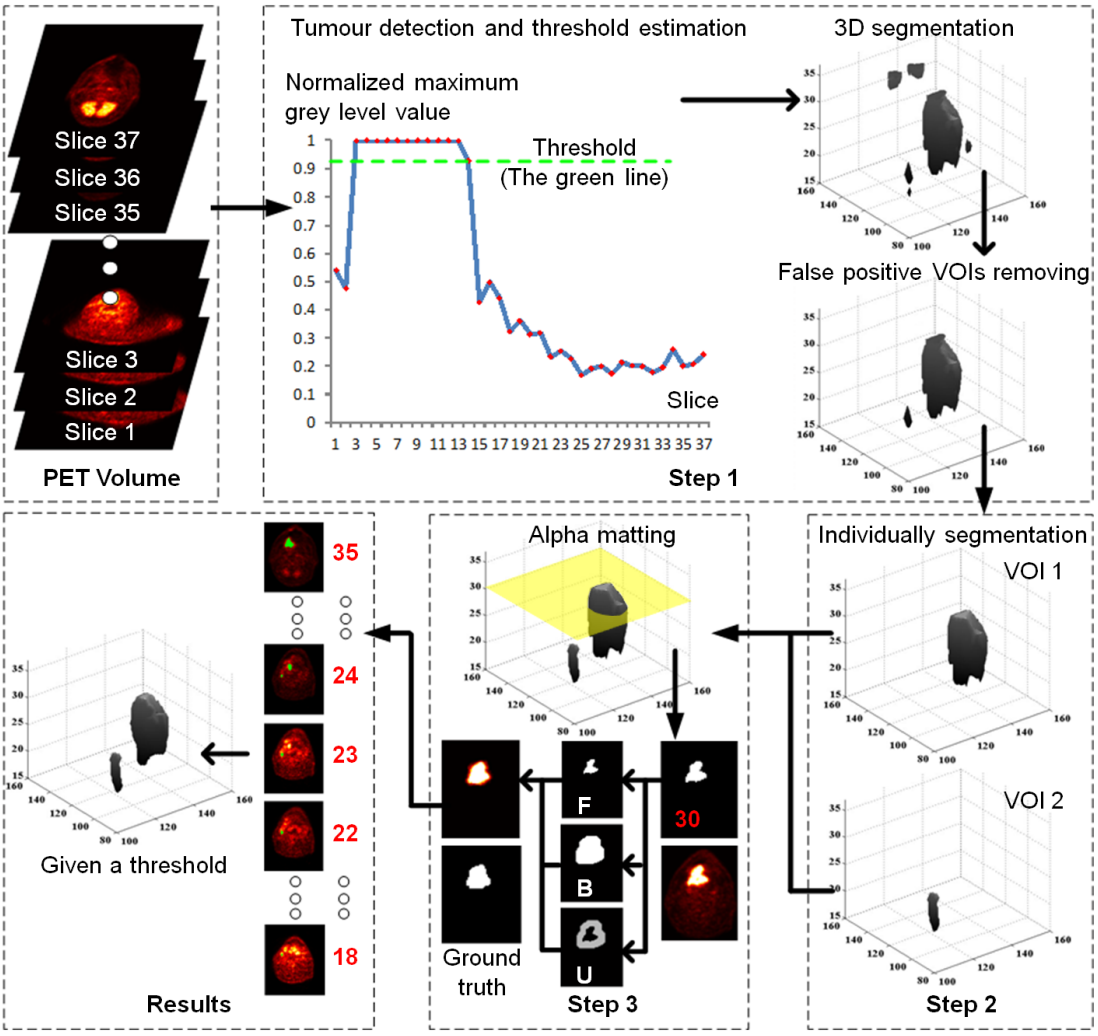


Figure 1: Segmentation results of the hierarchical segmentation approach. The tumour is shown in green on each slice of the original volume. Tumour images are labeled with red color numbers

2.1 Improved 3D Active Surface Modelling

A region-based active surface model [3] and a global convex segmentation model [5] are used to obtain an energy functional that can be minimized by convex optimization. Subsequently, the energy functional is minimized along with the deformation of the surface using a Split Bregman technique [6]. As in [10], the energy functional is formally defined as:

$$E(\phi, f_1, f_2) = \int \varepsilon_\varepsilon^x(\phi, f_1(x), f_2(x))dx + v\int \mid \nabla H_\varepsilon(\phi(x)) \mid dx + \mu P(\phi), \tag{1}$$

where ϕ denotes the surface of tumours, $P(\phi) = \int (1/2)(\mid \nabla\phi(x) \mid -1)^2 dx$ is the level set regularization term. In our work, we extend the method in [10] into 3D segmentation. The region-scalable energy functional is defined as

$$\varepsilon_\varepsilon^x(\phi, f_1(x), f_2(x)) = \sum_{i=1}^{2} \lambda_i \int K_\sigma(x-y) \mid f_i(x) - I(y) \mid^2 M_i(\phi(y))dy, \tag{2}$$

where K_σ is 3D Gaussian kernel functional. Given the locality of the kernel functional, the effect on ε generated by I(y) is almost zero when y is far away from x. The local fitting energy ε is determined by the value of σ. In Eq. 2, $M_1(\phi) = H(\phi), M_2(\phi) = 1 - H(\phi)$. The Heaviside functional H is usually approximated by a smooth functional $H_\varepsilon(x) = (1/2)\left[1 +\right.$

$(2/\pi)arctan(x/\xi)\big]$. As the grey level of tumours can vary between volumes, we change this functional to

$$H_\varepsilon(x) = \frac{1}{2}\left[1+\frac{2}{\pi}arctan\left(\frac{x-T}{\xi}\right)\right], \tag{3}$$

where T ($0 \leq T \leq 1$) is a normalisation value which is the tumour grey level threshold.

According to the derivation by Li et al. [3], the optimal functionals $f_1(x)$, $f_2(x)$ which minimize $E(\phi, f_1, f_2)$ are:

$$f_i(x) = \frac{K_\sigma(x) * [M_i(\phi(x))I(x)]}{K_\sigma(x) * M_i(\phi(x))}, \quad i = 1,2. \tag{4}$$

For fixed $f_1(x)$, $f_2(x)$, the functional ϕ can be written as:

$$\frac{\partial \phi}{\partial t} = \delta(\phi)\left[(-\lambda_1 e_1 + \lambda_2 e_2) + div\left(\frac{\nabla\phi}{\mid \nabla\phi \mid}\right)\right], \tag{5}$$

where δ is the derivative of H_ε. $e_i(x) = \int K_\sigma(y-x) \mid I(x) - f_i(y) \mid^2 dy, i = 1,2$. The simplified flow represents the gradient descent for minimizing the energy:

$$E(\phi) = \mid \nabla\phi \mid_1 + \langle \phi \cdot r \rangle, \tag{6}$$

where $r = \lambda_1 e_1 - \lambda_2 e_2$. Yang et al. [10] restricted the solution to lie in a finite interval in order to transform the constrained optimization problem to an unconstrained one. In this work, we constrain ϕ as $0 \leq \phi \leq 1$, which can guarantee a unique global minimal. The global convex model can be written as $min_{0\leq\phi\leq1}E(\phi) = min_{0\leq\phi\leq1}(\mid \nabla\phi \mid_1 + \langle \phi \cdot r \rangle)$. In [10], the minimization problem was written as $min_{0\leq\phi\leq1}E(\phi) = min_{0\leq\phi\leq1}(\int g \mid \nabla\phi \mid + \langle \phi \cdot r \rangle)$, where $g(\Theta) = 1/(1+\beta \mid \Theta \mid^2)$, β is a constant value. Then the Split Bregman method is used to minimize the energy functional [10]. When the optimal ϕ is found, we can find the segmented region $\Omega^k = \{x : \phi^k(x) > 0.5\}$.

2.2 Hierarchical Segmentation Scheme

As preprocessing, significantly noisy images are discarded before 3D volume data is generated. In the first segmentation step, we extract a closed surface area by using a low threshold as the initial surface to segment the PET volume data. The threshold T in Eq. 3 is estimated by calculating the average of the maximum grey level values of some selected slices in the volume. Specifically, we find the maximum grey level value for each slice. then all the values are normalized between 0 to 1. Subsequently, we select the slice which contains the maximum value. This slice will be utilised to compare with its three left neighbouring slices (named left neighbouring window). If the difference between this slice and any of the slice in the left window is less than the threshold φ, move the current slice to its closest left slice and compare the slice with its corresponding left window. This propagation carries on until the difference is above the threshold φ. The selected slices are propagated in this way. The same process is also imposed to the right side. Finally, we calculate the average of the maximum value from all the selected slices as the threshold T. For robust segmentation, a reliability metric is used to remove false positive VOIs. Specifically, the 6-connected neighborhood voxels are labeled in the results. If the tumour only exists in one or two slices, then the corresponding 3D labelling will be removed from the results.

The second segmentation step tries to find the non-detected regions that are selected in the first step. Specifically, morphology is used to dilate each 3D VOI generated in the previous step. We then use the generated surface as the initial closed surface, and segment the VOIs again using the local grey level voxels. The threshold T in this step is estimated by computing the average grey level values of the voxels in each dilated VOI.

2.3 Segmentation by Using Alpha Matting

Due to the partial volume effect and limited image resolution, the segmentation results generated by the first two steps are not accurate enough. To further improve the segmentation accuracy, we introduce an alpha matting method [1] into our segmentation pipeline. Instead of generating binary segmentation labels, an alpha matting techniques can generate a fractional alpha values between 0 and 1 for these pixels, which can be viewed as accurate soft segmentation.

To use alpha matting, a trimap has to be generated at first, which separates the image into three regions: definite foreground F, definite background B, and the unknown region U. Our system automatically generates this trimap. Specifically, we use morphology to erode the previous segmentation result with a circular structuring element to obtain the foreground F, then the background B can be generated by dilating the binary segmentation. The unknown area can be generated by using the combined results. To solve the alpha matting problem, we use the approach proposed by Levin et al. [1]. Solving the matting problem leads to a soft segmentation of VOIs in PET images.

3 Experiment

To evaluate our method, we use 2 PET images of head-and-neck cancer patients and 2 of a custom-built tumour phantom [7, 8]. All images were acquired using a hybrid PET/CT scanner (GE Discovery) and the metabolic tracer [^{18}F]FDG. Ground truth was derived for each phantom VOI by thresholding the CT images and for each patient tumour by manual delineation.

The detailed segmentation process for the example of a patient tumour is shown in Fig. 1. The clinical data sets are $256 \times 256 \times 37$ voxels. In the first step, all the maximum grey level values are generated from the 37 slices in the volume. Subsequently, the selected slices from 2 to 15 are identified by using the proposed criteria ($\varphi = 0.12$), then the threshold T in Eq. 3 is automatically estimated as 0.916. In Eq. 2, the 3D Gaussian kernel σ is 3 and $\lambda_i = 10$. In Eq. 3, $\xi = 0.1$. The improved active surface modelling result is used to segment the image and four false positive labellings are automatically removed. In the second step, the threshold T is recalculated for each dilated VOI and two VOIs are segmented again. We can see some non-detected regions are identified in this step. In the third step, the alpha matting algorithm is used to refine the segmentation on each slice. To compare against the ground truth which is a binary image, we threshold the soft segmentation generated by matting at half of the maximum grey level value for each slice, which leads to the final binary segmentation of VOIs.

Fig. 2 shows a comparison with other methods on real and phantom PET data. Over-segmentation in the red region of the original image is apparent for the PGVF and MAP-MRF EM methods. This is partly explained by partial volume effects and the low spatial resolution of the images but could also be due to these slice-wise methods ignoring volumetric grey

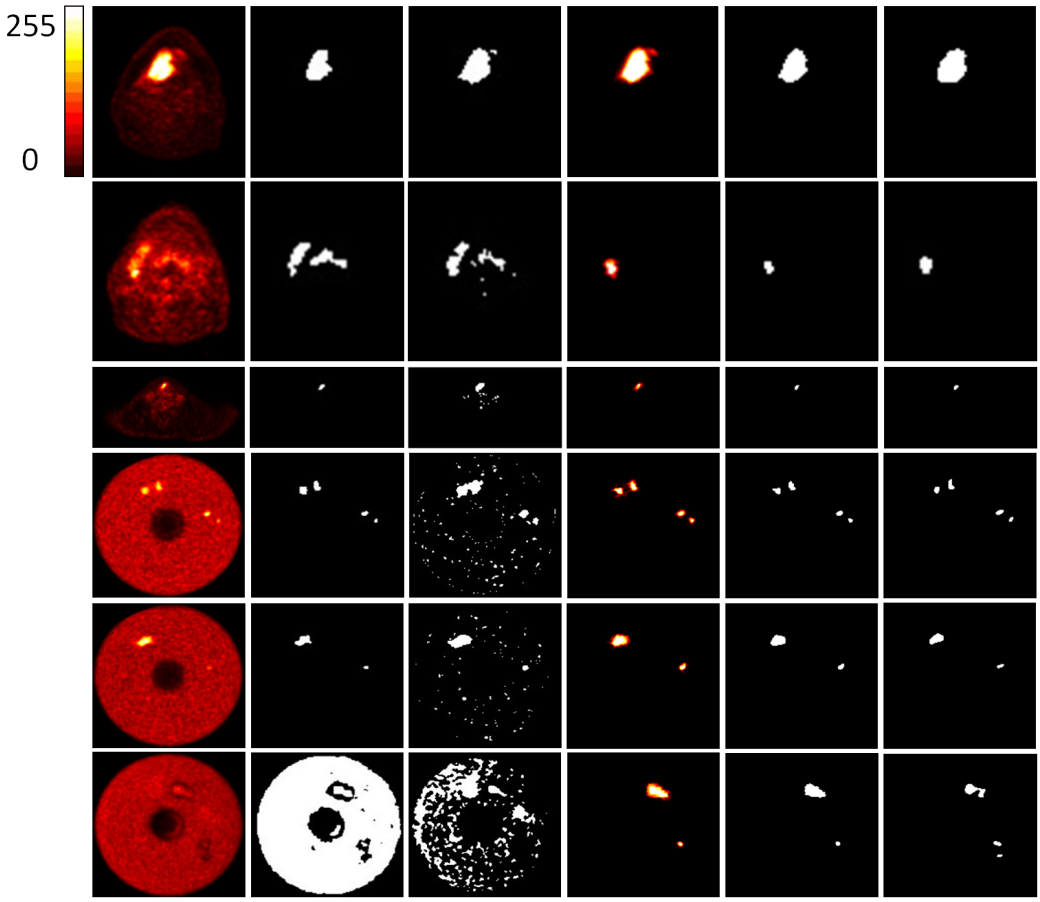


Figure 2: Segmentation of the patient VOIs. From left to right: Color bar shown from 0 to 255, PET data (patient data (slice 28, slice 21 in patient A, and slice 9 in patient B), phantom data (slice 22, slice 26 in phantom C, and slice 21 in phantom D)), PGVF, MAP-MRF EM, our method, our results after half threshold, ground truth.

level information. For the phantom data, the MAP-MRF EM method does not perform well on images with significant noise. The PGVF method fails to segment the abnormal region, because the density of the abnormal region is similar to the surrounding tissue. In contrast, the developed method performs well with these images and provides accurate segmentation results.

We evaluate segmentation accuracy in terms of overlap with ground-truth using the Dice similarity coefficient (DSC). We compare the DSC of our method and alternative approaches [2, 4]. Tab. 1 shows the mean DSC for real and phantom PET as well as the overall mean. Our results show improvements compared to previous methods.

4 Discussion and Conclusions

This paper presents a novel PET image segmentation scheme, based on an improved region-based active surface modelling method and alpha matting. Our method has the following advantages. First, the 3D voxel information is considered in a hierarchical scheme, which can largely eliminate segmentation errors. Second, the alpha matting technique is introduced to PET image segmentation for the first time, which can effectively deal with the partial volume effect problem. Finally, compared with previous approaches, our method is more robust on PET imaging segmentation. As future work, we will further evaluate our method

Table 1: Dice index for tumour segmentation on PET data

Algorithm	PGVF [2]	MRF [4]	Our method (half thresholding results)
Patient images	0.562	0.430	0.706
Phantom images	0.423	0.383	0.664
All images	0.493	0.407	0.685
Standard deviations	0.234	0.218	0.142

on larger clinical datasets.

References

[1] A.Levin, D.Lischinski, and Y.Weiss. A closed form solution to natural image matting. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 30:228-242, 2008.

[2] C.Hsu, C.Liu, and C.Chen. Automatic segmentation of liver PET images. *Computerized Medical Imaging and Graphics*, 32:601-610, 2008.

[3] C.Li, C.Kao, C.John, and Z.Ding. Minimization of Region-Scalable Fitting Energy for Image Segmentation. *IEEE Trans. Image Processing*, 17, 1940-1949, 2008.

[4] H.Gribben, P.Miller, H.Wang, K.Carson, A.Hounsell, and A.Zatari. Automated MAP-MRF EM labelling for volume determination in PET. *ISBI*, 1-4, 2008.

[5] T.F.Chan, S.Esedoglu, and M.Nikolova. Algorithms for finding global minimizers of denoising and segmentation models. *SIAM Journal on Applied Mathematics*, 66:1632-1648, 2006.

[6] T.Goldstein, X.Bresson, and S.Osher. Geometric Applications of the Split Bregman Method: Segmentation and Surface Reconstruction. *Journal of Scientific Computing*, 45:272-293, 2010.

[7] T.Shepherd, M.Teräs, and H.Sipilä. New Physical Tumour Phantom and Data Analysis Technique Exploiting Hybrid Imaging and Partial Volume Effects for Segmentation Evaluation in Radiation Oncology. *European Journal of Nuclear Medicine and Molecular Imaging*, 37:S221, 2011.

[8] T.Shepherd, M.Teräs, and H.Sipilä. Results of the Contouring Challenge, Software Session II at the XII Turku PET Symposium, Turku PET Centre, Finland. *[online]. Available: http://www.turkupetcentre.net/PET_symposium_XII_software_session /ContouringChallengeResults/*, 2011.

[9] Y.E.Erdi, O.Mawlawi, S.M.Larson, M.Imbriaco, H.Yeung, R.Finn, and J.L.Humm. Segmentation of lung lesion volume by adaptive positrone mission tomography image thresholding. *Cancer*, 80:2505-2509, 1997.

[10] Y.Yang, C.Li, C.Kao, and S.Osher. Split Bregman method for minimization of region-scalable fitting energy for image segmentation. *Lecture Notes in Computer Science*, 6454, 117-128, 2010.

Level Set based Tracking for Cell Cycle Analysis using Dynamical Shape Prior

Yan Nei Law
lawyn@bii.a-star.edu.sg
Hwee Kuan Lee
leehk@bii.a-star.edu.sg

Bioinformatics Institute
30 Biopolis Street, #07-01 Matrix,
Singapore 138671

Abstract

Automated cell tracking in populations is very crucial for studying dynamic cell cycle behaviors. However, high accuracy of each step is essential to avoid error propagation. In this paper, we propose an integrated three-component system to tackle this problem. We first model the temporal dynamics of shape change using an autoregressive model, which is used for estimating the shape and the location of the current object. We then segment the cell using active contour model starting from the predicted shape. Finally, we identify its phase using Markov model. This information is also used for further fine-tuning the segmentation result. We applied this approach for tracking HeLa H2B-GFP cells and high accuracy validation results confirm the usefulness of our integrating approach.

1 Introduction

To study dynamic cell cycle behaviors [6], single cell tracking in populations is a crucial step to obtain quantitative measurements. In many studies, this process was carried manually which is very time-consuming and the results could be subjective. To automate this process, we introduce a new system for individual cell tracking in this paper. The system shown in Fig. 1 consists of three components: (C1) Learning shape prior for tracking, (C2) Segmentation using prior, and (C3) Cell phase identification. In contrast with most of the existing methods which focused on these subproblems separately, this approach makes use of available biological prior knowledge for improving the performance of the whole process.

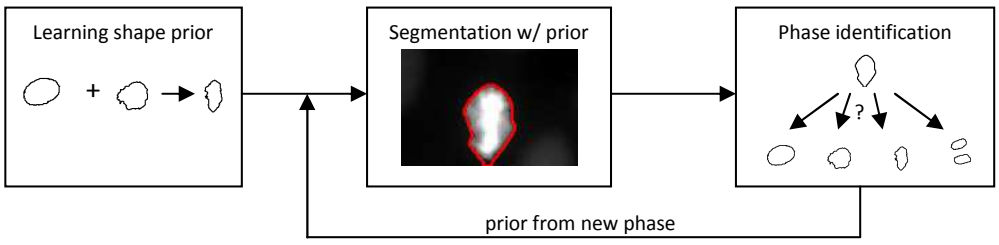


Figure 1: Flowchart of the proposed system.

For (C1), we model the temporal dynamics of shape change using an autoregressive model [1, 4, 7]. To obtain a more reliable model, we use a low-dimensional formulation

to represent a shape. The model is used for estimating the shape and the location of the current object. Details will be present in Section 2. For (C2), we use initialization from (C1) and segment the cell using an active contour model [5]. This step greatly improves the segmentation accuracy. Details will be present in Section 3. For (C3), we model this using Markov process and classify the segmented cell. The segmentation result is further enhanced based on a given sample shape if the cell undergoes a phase transition. Details will be present in Section 4. Note that our approach is very generic and can be applied for a large variety of cell images by using different training examples and sample shape examples.

2 Learning Shape Prior for Tracking

2.1 Finite-dimensional curve representation

Given a sequence of N training shapes, we define the centroid as the center of mass and the orientation as the angle between the horizonal axis and the major axis of the ellipse that has the same second-moments as the shape. After a suitable transformation, all the shapes are centered and have 0° orientation. Their signed distance functions (SDFs) $\{\phi_1, \dots \phi_N\}$ are uniquely determined [4]. The mean shape ϕ_0, is computed by taking the mean of the SDFs, $\phi_0 = \frac{1}{N}\sum \phi_i$. The n largest eigenmodes $\psi = (\psi_1, \dots \psi_n)$ with $n \ll N$ are computed by Principal Component Analysis (PCA) [2]. Then each training shape can be approximated by a linear combination of the eigenmodes:

$$\phi_i(x) \approx \phi_0(x) + \sum_{j=1}^{n} \alpha_{ij}\psi_j(x),$$

where $\alpha_{ij} = \langle \phi_i - \phi_0, \psi_j \rangle \equiv \int (\phi_i(x) - \phi_0(x))\psi_j(x)\, dx$. This representation has been widely used for constructing shape priors [1, 4]. Given an arbitrary shape ϕ (after a suitable transformation), it can be approximated by a shape vector of the form $\alpha_\phi = (\langle \phi - \phi_0, \psi_1 \rangle, \dots, \langle \phi - \phi_0, \psi_n \rangle)$. For the transformation parameters (translation and rotation), instead of using the absolute transformation θ, we use the incremental transformation $\Delta\theta$, i.e., difference between the current transformation and the previous transformation. Together with the shape parameters, a sequence of shapes can be represented by a sequence of combined vectors

$$\mathbf{v}_i = \begin{pmatrix} \alpha_i \\ \Delta\theta_i \end{pmatrix}. \tag{1}$$

The main advantages of this representation are its simplicity and robustness.

2.2 Learning temporal dynamics of shape change

Given the shape and transformation parameters $\{\mathbf{v}_1, \dots, \mathbf{v}_{t-1}\}$ obtained from the previous shapes $\{\phi_1, \dots, \phi_{t-1}\}$, we aim to estimate the parameters of the shape at time t. To achieve this, we model this process using an autoregressive model (AR) of order 2 [7], i.e.,

$$\mathbf{v}_t = \mu + A_1\mathbf{v}_{t-1} + A_2\mathbf{v}_{t-2} + \eta, \tag{2}$$

where μ is the mean and η is zero-mean Gaussian noise with covariance Σ. To estimate the model parameters, we use a stepwise least squares method proposed in [7]. We then predict the shape vector $\mathbf{v}_t$ as $\tilde{\mathbf{v}}_t = \mu + A_1\mathbf{v}_{t-1} + A_2\mathbf{v}_{t-2}$. Since the estimation provides information about the shape and the location of the current object, this serves as a purpose of tracking embedded into the system. Figure 2 shows the approximation of a HeLa cell in each phase.

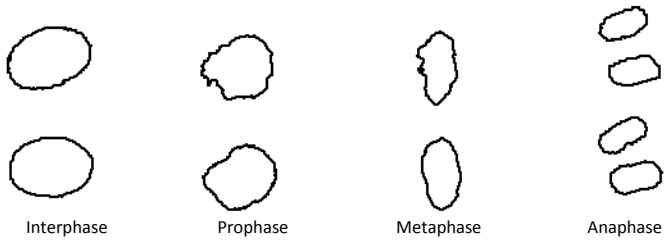


Figure 2: The approximation (below) of a cell (above) by the first six principal components.

3 Segmentation with shape prior initialization

Given an image $I_t : \Omega \to \mathbf{R}$ and a sequence of previous segmented images, our aim is to track the object in I_t making use of the information derived from the estimated shape vector $\tilde{\mathbf{v}}_t$, which is often very accurate. In image processing applications, many effective active contour models [8] require a good initialization. The estimated shape $\tilde{\mathbf{v}}_t$ we generated can serve this purpose well.

To segment the object, we choose the edge-based active contour model proposed in [5] among many segmentation methods [8] because it inherits some advantages of level set methods such as simple representation for contours of complex topology and simple numerical computations on regular grids, while it has an intrinsic capability of maintaining regularity of the level set function. The objective is to find a level set function $\phi : \Omega \to \mathbf{R}$ such that the following energy functional $F(\phi)$ is locally minimized:

$$F(\phi) = a\int_\Omega p(|\nabla\phi|)\,dx + b\int_\Omega g\delta(\phi)|\nabla\phi|\,dx + c\int_\Omega gH(-\phi)\,dx \tag{3}$$

where $a, b > 0$ and $c \in \mathbf{R}$ are the coefficients of the three terms, p is a potential function, g is an edge indicator function, δ and H are the Dirac delta function and the Heaviside function respectively. Here, we use

$$p(s) = \begin{cases} \frac{1}{(2\pi)^2}(1-\cos(2\pi s)), & \text{if } s \le 1 \\ \frac{1}{2}(s-1)^2, & \text{if } s \ge 1, \end{cases} \qquad g = \frac{1}{1+|\nabla G_\sigma * I_t|^2}$$

where G_σ is a Gaussian kernel with a standard deviation σ. To minimize Eq. (3), we apply gradient descent method. We denote by $\frac{\partial F}{\partial \phi}$ the (Gâteaux) derivative of F with respect to ϕ. Then, a single step of the gradient descent method computes the next iterate by the update $\phi^{k+1} = \phi^k - \Delta t \frac{\partial F}{\partial \phi}(\phi^k)$, where $\Delta t > 0$ is the step length parameter. The detailed numerical implementation can be found in [5]. Presumably, the contour stops at the object's boundary.

4 Cell phase identification

In this section, we identify the cell phase S_t based on the segmentation result obtained by the previous step. Here, we model this as a maximum-a-posteriori (MAP) problem. Such an approach has been widely used for modeling state transition [9, 10]. Given the current image I_t and the previous phases $S_{1:t-1} = \{S_1, \ldots, S_{t-1}\}$, we denote by $P(S_t|I_t, S_{1:t-1})$ the posterior probability of the current phase S_t. Then its maximizer S_t^* will be the optimal current phase. Based on the Bayesian formula, the expression can be rewritten as

$$P(S_t|I_t, S_{1:t-1}) = \frac{P(I_t|S_t, S_{1:t-1})P(S_t|S_{1:t-1})}{P(I_t|S_{1:t-1})} \propto P(S_t|S_{t-1})P(I_t|S_t).$$

The last expression is based on the assumptions that I_t and $S_{1:t-1}$ are uncorrelated and S_t depends only on the previous phase S_{t-1}. For simplicity, we consider the "left-right" model, i.e., transitions could happen only between consecutive phases. Also, we track the cell in one cycle. Moreover, the probabilities of all possible transitions are equal. These assumptions could be expressed using the following matrix:

$$\begin{array}{c|cccc} & I & P & M & A \\ I & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ P & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ M & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ A & 0 & 0 & 0 & 1 \end{array}$$

where I,P,M,A represent interphase, prophase, metaphase and anaphase, respectively. For the likelihood function $P(I_t|S_t)$, we use the following similarity measure between the segmented shape ϕ and the sample shape ϕ_S associated with S where $S = \{I,P,M,A\}$:

$$P(I_t|S_t) \propto D(\phi,\phi_S) = \exp(-\sum_i |f^i_\phi - f^i_{\phi_S}|),$$

where f_ϕ is the shape and context features. Here, we use eight features suggested in [10]: maximum intensity, minimum intensity, standard deviation of intensity, mean intensity, length of major axis, length of minor axis, perimeter and compactness (perimeter^2/(4$\pi\cdot$ Area)).

The phase information can be used to enhance the segmentation result. If the cell undergoes a phase transition, we repeat the segmentation step with the sample shape of the new (optimal) phase as an initial guess. The sample shape is first transformed to obtain the same centroid and orientation with the segmented shape. With this biological prior knowledge, the segmentation result could be further improved by this additional step. This will be illustrated in the experimental section.

We then use the final segmented shape to compute the shape vector $\mathbf{v}_t$ and use it for predicting the shape prior for the next image.

5 Experimental results

In this section, we evaluate the performance of the proposed method empirically in various aspects: 1) temporal sampling rate, 2) system combination and 3) dwelling time.

The cells used in this paper are HeLa H2B-GFP cell line from [3] and [10] under different conditions. Frames were acquired at 3-minute intervals and 15-minute intervals respectively, with a time-lapse fluorescence microscopy. Each image is of 97×130 pixels.

To estimate the parameters of the model in Eq. (2), we manually segmented a sequence of 66 training images of a cell nucleus and computed the corresponding shape vectors and the eigenmodes ($n = 6$) described in Section 2. For the initial setting $\mathbf{v}_1$ and $\mathbf{v}_2$ in Eq. (2), we first manually segment the first two images in the test sequence and compute their shape vectors $\mathbf{v}_1$ and $\mathbf{v}_2$.

The percentage intersection between the manual and the computed segmentation is used as a measure of segmentation accuracy and the mean accuracy with error bar is reported. The error bar indicates the standard deviation of the accuracy.

5.1 Sampling rate

In this test, we use a sequence of 40 images of cell nuclei as the test set. From this sequence, we generate four other sequences by duplicating or skipping images to simulate sequences

with higher or lower sampling rate respectively. Figure 3A shows the mean accuracy of the segmentation results vs. the temporal sampling rate. Observe that the accuracy remains high as the sampling rate is higher than 1. However, it degrades slightly when the rate is decreased. It shows that it is difficult to capture a dramatic change in shape due to low sampling rate stream. But the overall accuracy is still high. Thus, our algorithm is robust.

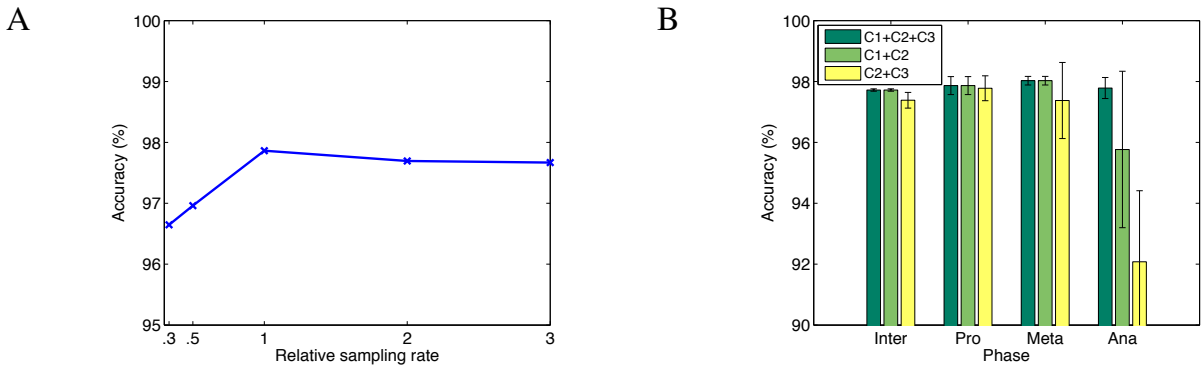


Figure 3: Mean accuracy (A) vs. sampling rates (B) with different system combinations of (C1) learning prior, (C2) segmentation and (C3) phase identification.

5.2 System combination

To evaluate the usefulness of components (C1) learning prior and (C3) phase identification, we segment the above image sequence using different system combinations. Figure 3B shows the mean accuracy of the segmentation results in different phases. For the combination of (C2)+(C3), we use the final segmentation from one frame as the initialization of the next. Observe that the accuracies in the first three phases are comparable. However, the accuracy in the anaphase is degraded when either the step (C1) or (C3) is absent. Note that the morphological change from metaphase to anaphase is dramatic. Both the initialization learned from the dynamical model and that assigned by phase identification improve significantly the results.

5.3 Image sequence with short dwelling time of prophase

In this test, we use two other sequences of 20 images of cell nuclei having different sampling rates (A: low, B: high). For these two sets, the cells have a very short dwelling time of prophase which makes phase identification more difficult. Figure 4 shows the mean accuracy and some resulting images from the two sets in different phases. Observe that a short dwelling time of prophase causes a miss of metaphase identification. The error also propagates to the next phase. The effect is shown in the result of set A. However, due of its higher sampling rate, there are more frames sampled in metaphase in set B and hence is able to reduce the effect. Overall, the accuracy in anaphase is lower than that in the other phases.

6 Conclusion

In this paper, we introduced a three-component system for individual cell tracking in populations. This approach makes use of available biological prior knowledge to improve the performance of the process. On the other hand, it is very generic and can be applied for a large variety of cell images by using different training examples and sample shape examples. While we only tested the method on tracking single cells, it is natural to generalize the system by using several level set functions for multiple cell tracking. A promising future work is to optimize the performance of the integration approach to make it more practical.

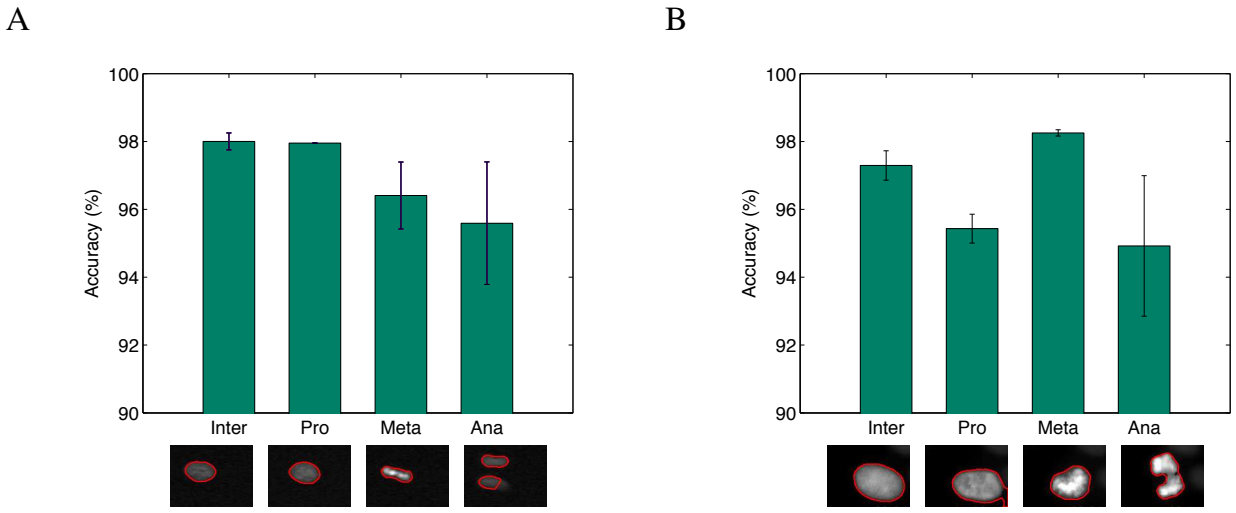


Figure 4: Mean accuracy and some resulting images in different phases.

Acknowledgement

This work was supported (in part) by the Biomedical Research Council of A*STAR (Agency for Science, Technology and Research), Singapore.

References

[1] D. Cremers. Dynamical statistical shape priors for level set-based tracking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(8):1262–1273, 2006.

[2] R. Duda, P. Hart, and D. Stock. *Pattern Classification*. John Wiley & Sons, NY, 2001.

[3] S. Elowe, K. Dulla, A. Uldschmid, X. Li, Z. Dou, and E.A. Nigg. Uncoupling of the spindle-checkpoint and chromosome-congression functions of BubR1. *J. Cell Sci.*, 123.

[4] M. Leventon, W. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Proc. CVPR*, pages 316–323, 2000.

[5] C. Li, C. Xu, C. Gui, and M.Fox. Distance regularized level set evolution and its application to image segmentation. *IEEE Trans. Image Process.*, 19(12):3243–3254, 2010.

[6] H. Lodish, A. Berk, S. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell. *Molecular Cell Biology*, chapter 13. W.H. Freeman & Co., 2000.

[7] A. Neumaier and T. Schneider. Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans. Math. Software*, 27(1):27–57, 2001.

[8] O. Scherzer. *Handbook of Mathemtical Methods in Imaging*. Springer, 2011.

[9] X. Zhou, X. Li, and W. Hu. Level set tracking with dynamical shape priors. In *Proc. ICIP*, pages 1540–1543, 2008.

[10] X. Zhou, F. Li, J. Yan, and S. Wong. A novel cell segmentation method and cell phase identification using markov model. *IEEE Trans. Info. Technol. Biomed.*, 13(2):152–157, 2009.

Graph-based Segmentation of Optimal IVUS Media-Adventitia Border using Shape Prior

Ehab Essa[1]
csehab@swansea.ac.uk
Xianghua Xie[1]
X.Xie@swansea.ac.uk
Igor Sazonov[2]
I.Sazonov@swansea.ac.uk
Perumal Nithiarasu[2]
P.Nithiarasu@swansea.ac.uk
Dave Smith[3]
Dave.Smith@abm-tr.wales.nhs.uk

[1] Department of Computer Science
Swansea University, UK
[2] College of Engineering
Swansea University, UK
[3] ABM University NHS Trust
Swansea, UK

Abstract

We present a shape prior based graph cut method which does not require user initialisation. The shape prior is generalised from multiple training shapes, rather than using singular templates as priors. Weighted directed graph construction is used to impose geometrical and smooth constraints learned from priors. The proposed cost function is built upon combining selective feature extractors. A SVM classifier is used to determine an optimal combination of features in the presence of various IVUS tissues and artefacts. Comparative analysis on manually labelled ground-truth shows superior performance of the proposed method compared to conventional graph cut methods.

1 Introduction

Intra-vascular Ultrasound (IVUS) imaging is a catheter-based technology, which assesses the severity and morphology of the coronary artery stenosis. The media-adventitia border represents the outer coronary arterial wall located between the media and adventitia. The media layer exhibits as a thin dark layer in ultrasound and has no distinctive feature. It is surrounded by fibrous connective tissues called adventitia. The media-adventitia border in IVUS is disrupted by various forms of artefacts such as acoustic shadow or reverberation which can be caused by catheter guide-wire or fatty and cholesterol materials deposit and fibrosis formed inside the artery. Fig. 1 gives an example of IVUS image.

Common approaches to IVUS segmentation include graph cut and deformable modelling, and usually requires user initialisations [3, 7, 12]. The use of shape prior has shown to be a promising approach to tackle the ambiguities in identifying media-adventitia border. For example, Unal *et al.* [13] used principal component analysis (PCA) to generalise the shape variation. The initialisation of the media-adventitia border is based on the maximum

Figure 1: Pre-Processing steps. (a) Original IVUS image. (b) Polar transformed image. (c) After removing the catheter region.

intensities which can be distracted by fibrotic tissue and other imaging artefacts especially in high frequency ultrasound images.

The integration of shape prior into graph-based methods is still a challenging problem. Freedman and Zhang [6] defined the shape template as a distance function and embedded the average distance between every pair of pixels into the neighbourhood edges in the graph. However, this method effectively requires the user to place landmarks to define the initial shape. In [9], the authors proposed an iterative graph cut method. Kernel PCA was used to build the shape model. Affine transformation of shapes was not taken into account, and user initialisation is necessary.

In this paper, we propose an efficient graph cut algorithm to segment media-adventitia border in IVUS images without user initialisation. Prior knowledge of the IVUS tissue type is used to optimise the cost function. Weighted directed graph is defined to impose geometric constraints. The proposed method can handle large variations in training shapes which may undergo affine transformations. The generalised shape prior is incorporated in the cost function, as well as in the graph construction.

2 Proposed Method

The images are first transformed from Cartesian coordinates to polar coordinates and the catheter regions are removed (see Fig. 1). This transformation not only facilitates our feature extraction and classification but also transfers a closed contour segmentation to a "height-field" segmentation. The border to be extracted intersects once and once only with each column of pixels. This particular form of segmentation allows us to construct a node-weighted directed graph, on which a minimum path can be found without any user initialisation.

2.1 Feature extraction and classification

The imaging artifacts generally have large responses to image gradient based feature extraction. In this work, we propose to detect those artifacts and treat them differently when incorporating into the cost function. To highlight the media-adventitia border, we use a combination of derivative of Gaussian (DoG) features and local phase features. A set of first and second order DoG filters are applied to enhance the intensity difference between media and adventitia. Local phase [10] has shown to be effective in suppressing speckles in ultrasound images. We use the dark symmetry feature to highlight bar-like image patterns, which are useful to detect the thin media layer.

For those parts of media-adventitia border that are beneath various forms of image artifacts, such as calcification, their image features are suppressed by those artifacts. Hence, it is desirable to detect those artifacts and treat those columns of pixels differently to others.

Instead of a usual attempt of localising those image artifacts based on intensity profile, e.g. [5, 13], which is problematic, we classify entire columns of pixels that contain those image artifacts. The detection result will then have an influence on the formulation of the cost function. To this end, we train a SVM classifier to classify individual columns of pixels in the polar coordinates into one of the following five categories: calcification, fibrous plaque, stent, guide-wire artifact, and normal tissue or soft plaque.

2.2 Boundary based cost function

The boundary based energy term can be expressed as $E_B = \sum_{V \in S} \hat{c}_B(x,y)$, where $\hat{c}_B$ denotes the normalised cost function ($\hat{c}_B(x,y) \in [0,1]$) and S is a path in the directed graph. The formulation of the pre-normalisation cost function, c_B, is determined by the SVM classification result as presented below.

For normal tissue (or soft plaque), the media layer has a good contrast to adventitia. Hence, c_B is defined as $c_B(x,y) = D_1(x,y) - D_2(x,y)$ where D_1 is a summation of raw filtering response of the first order DoG at four different orientations and D_2 denotes maximum response of second order DoG filtering from different orientations across three scales. invariant measurement of bar-like feature.

Calcified plaque exhibits strong edge features and casts varying degree of acoustic shadow. Thus, we use the second order DoG responses to suppress calcification and enhance possible media layer. Fibrous tissue behaves similarly to calcification, except in majority cases media-adventitia border is still discernible. Hence, bar feature detection is more appropriate and to enhance the effect we combine it with phase symmetry feature, i.e. $c_B(x,y) = -D_2(x,y) - FS(x,y)$ where FS is the local phase feature.

The cost for stent case is the second order DoG responses with more weight underneath the stent region. For guide-wire artifact, we do not extract any feature and a positive constant is used as their cost value, since it is a small area and cast acoustic shadowing.

2.3 Shape prior based cost function

The energy term for shape prior can be expressed as:

$$E_S = \sum_{V \in S} c_S(x,y) + \sum_{(p,q) \in \mathcal{N}} f_{p,q}(S(p) - S(q)), \tag{1}$$

where c_S denotes the cost function associated to prior and f is a convex function penalizing abrupt changes in S between neighbouring columns p and q in the set $\mathcal{N}$ of neighbouring columns in the graph. The second term is realized through graph construction, detailed in the following Section 2.4.

Each shape in the training set is treated as a binary template, ψ where the area inside shape is one and the outside area is zero. The distance between two templates ψ^a and ψ^b is defined using a discrete version of Zhu and Chan distance [4]: $d^2(\psi^a, \psi^b) = \sum_P (\psi^a - \psi^b)^2$, where P denotes the image domain. This distance measure is a true metric and is not influenced by image size. Affine transformation is also applied to remove the influence from geometrical transformation. Let $\Psi = \psi^1, \ldots, \psi^N$ denote the N number of aligned shapes from the training set. Given a possible cut in the graph which produces an aligned binary shape f, its similarity to a shape template ψ^n in the training set is computed as $\alpha(f, \psi^n) = exp(-\frac{1}{2\sigma^2} d^2(f, \psi^n))$. Thus, the likelihood of this particular cut can be evaluated by taking into account of all training shapes: $\omega_s = \sum_{n=1}^{N} \alpha(f, \psi^n)\psi^n / \sum_{n=1}^{N} \alpha(f, \psi^n)$. The labelling of the shape likelihood and initial cut needs to be compared in order to assign appropriate

terminal arcs. For simplicity and efficiency, we search the lower bound of the likelihood function in the image domain and assign negative cost to this lower bound whileas other regions are assigned positive cost to form our shape cost term, c_S. Hence, we eliminate the need to identify the terminal connection type.

2.4 Graph construction using shape prior

Conventional graph cut, such as [1], generally requires user initialisation. In [8], Li *et al.* introduced a graph construction method which transforms surface segmentation in 3D into computing a minimum closed set in a directed graph without any user initialisation. Very recently, Song *et al.* [11] extended the method to incorporate shape prior by assuming that the relationship between every neighbouring columns can be represented by a parametric normal distribution. In this work, we propose an extension of this method by performing affine transformation based alignment of the training shapes and the use of a shape similarity metric in computing the inter-column parametric distribution. Additonally, we alter the inter-column arcs to permit more influence from the boundary features.

Let $G = \langle V, E \rangle$ denote the graph, where each node $V(x,y)$ corresponds to a pixel in the transformed IVUS image $I(x,y)$ in polar coordinates. The graph G consists of two arc types: intra-column arcs and inter-column arcs. For intra-column, along each column every node $V(x,y)$, where $y > 0$, has a directed arc to the node $V(x,y-1)$ with $+\infty$ weight assigned to the arc to ensure that the desired interface intersects with each column exactly once. Inter-column acts as a hard constraint and derived from shape prior. In calculating the shape prior cost function, an initial cut is first obtained by minimising the boundary based cost function alone. Note, this is fully automatic and there is no need for user initialisation. The training shapes are then affine-aligned to our initial graph cut. The similarity measurement α is used to choose a set of closest templates. The inter-column changes are then generalised using mean $m_{p,q}$ and standard deviation $\sigma_{p,q}$ at individual column. These statistics are then used in determining maximum and minimum distances when connecting neighbouring columns in graph construction, i.e. $\bar{\Delta}_{p,q} = m_{p,q} + c \cdot \sigma_{p,q}$, $\underline{\Delta}_{p,q} = m_{p,q} - c \cdot \sigma_{p,q}$, and c is a real constant.

Additional inter-column arcs are necessary in order to allow smooth transition. That is intermediate values, $h \in [\underline{\Delta}_{p,q}, \bar{\Delta}_{p,q}]$, are used to construct inter-column arcs. The direction of these arcs is based on the first order derivative of the function $f_{p,q}(h)$ as in (1). Here, we employ a quadratic function, $f_{p,q} = \lambda(x - m_{p,q})^2$, where λ is a weighting factor for smoothness constraint. If $f'_{p,q}(h) \geq 0$ an arc from $V(x,y)$ to $V(x+1, y-h)$ is established; otherwise, the arc is connected from $V(x+1,y)$ to $V(x,y+h)$. The weight for these arcs is assigned as the second order derivative of $f_{p,q}$ for the intermediate values of h, and $+\infty$ weight for the endpoints of the interval. Note, when $f'_{p,q}(h) = 0$, only single arc is defined to reduce the shape prior influence in presence of strong boundary features, instead of using bi-directional arcs on the mean difference $m_{p,q}$.

2.5 Compute the minimum closed set

The cost function $C(x,y) = c_B(x,y) + c_S(x,y)$ is inversely related to the likelihood that the border of interest passes through pixel (x,y). The weight for each node on the directed graph can be assigned as:

$$w(x,y) = \begin{cases} C(x,y) & \text{if } y = 0, \\ C(x,y) - C(x,y-1) & \text{otherwise.} \end{cases} \tag{2}$$

Segmenting the media-adventitia is equivalent to finding the minimum closed set in the directed graph. The *s-t* cut algorithm [2] can then be used to find the minimum closed

Table 1: Quantitative comparison to $s-t$ cut. AD: area difference in percentage; AMD: absolute mean difference in pixel in comparison to groundtruth.

	$s-t$ cut		proposed method	
	AD	AMD	AD	AMD
Mean	22.54	23.91	8.71	9.46
Std.	8.87	7.49	4.99	5.36

set, based on the fact that the weight can be used as the base for dividing the nodes into nonnegative and negative sets. The source s is connected to each negative node and every nonnegative node is connected to the sink t, both through a directed arc that carries the absolute value of the cost node itself. The segmented media-adventitia may still contain local oscillations. Here, efficient 1D RBF interpolation using thin plate base function is used to obtain the final interface.

3 Experimental results

A total of 1197 IVUS images of 240×1507 pixels in the polar coordinates from 4 pullbacks acquired by a 40 MHz transducer are used to evaluate the proposed method. These images contain various forms of fibrous plaque, calcification, stent, and acoustic shadow. For all the tested images, ground-truth via manual labelling is available for quantitative analysis. The training shape set contains 278 images. Another set of 138 images are used to train SVM classifier.

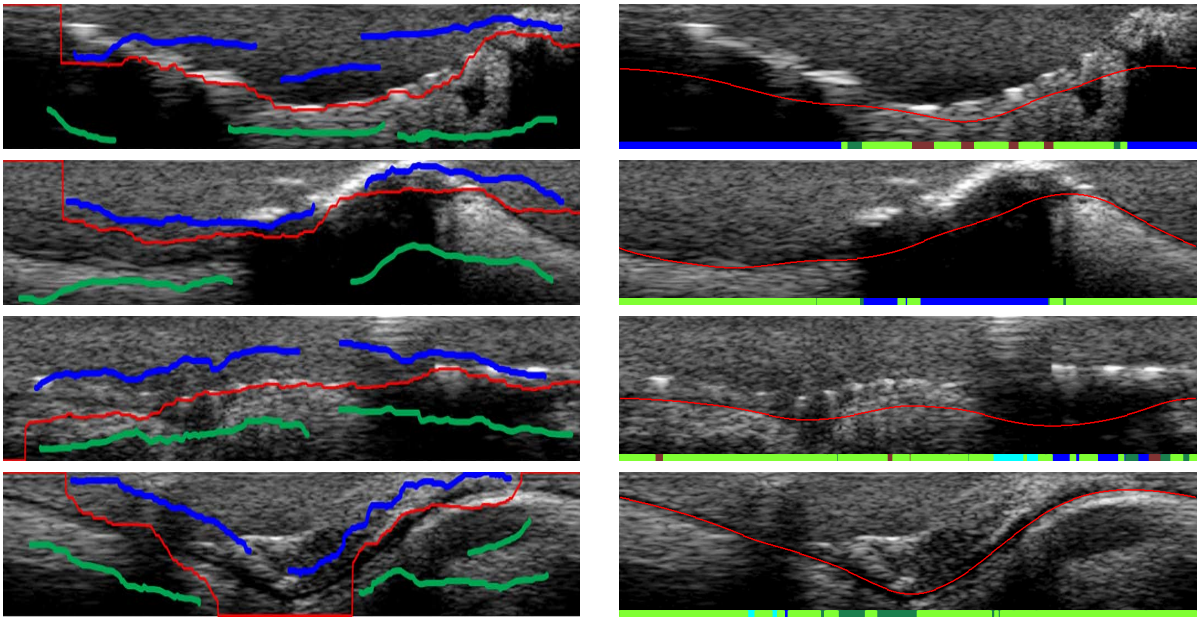
First, we compared our method against the $s-t$ cut algorithm [1]. Despite reasonable care in initialisation as shown Fig. 2(a), the $s-t$ cut result was not satisfactory. The corresponding results of the proposed method are shown in the second column. The bottom of the each image shows the classification result of detecting different types of tissue. The proposed method achieved better accuracy and consistency. The quantitative comparison was carried out on a randomly selected subset of 50 images, since manual initialisation of 1197 images is too labour intensive. Table 1 shows that the proposed method clearly outperformed $s-t$ cut in both area difference measure (AD) and absolute mean difference measure (AMD) based on groundtruth. Next, the proposed method was tested on the full dataset (1197 images) and its performance based on labelled groundtruth can be summarised as: 7.88% mean AD with standard deviation of 5.79 and 8.49 pixel mean AMD with standard deviation of 5.91. This is marginally better than the first subset. Fig. 3 shows example comparisons to groundtruth.

4 Conclusions

We presented an automatic graph based segmentation method for delineating the media-adventitia border in IVUS images. Boundary based features were dynamically selected to optimise the cost function. The use of multiple training shapes proved to be beneficial. The generalised shape prior was used in both incorporating the cost function but also graph construction. Qualitative and quantitative results on a large number of IVUS images showed superior performance of the method.

References

[1] Y. Boykov and G. Funka-Lea. Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. *IJCV*, 70(2):109–131, 2006.



(a) (b)

Figure 2: (a) $s-t$ cut result (red) with user initialization (object: blue, background: green). (b) proposed method result; the bottom of each image also shows the classification result: calcified plaque (blue), fibrotic plaque (dark green), stent (dark red), guide-wire shadowing (cyan), and soft plaque/normal tissue (light green).

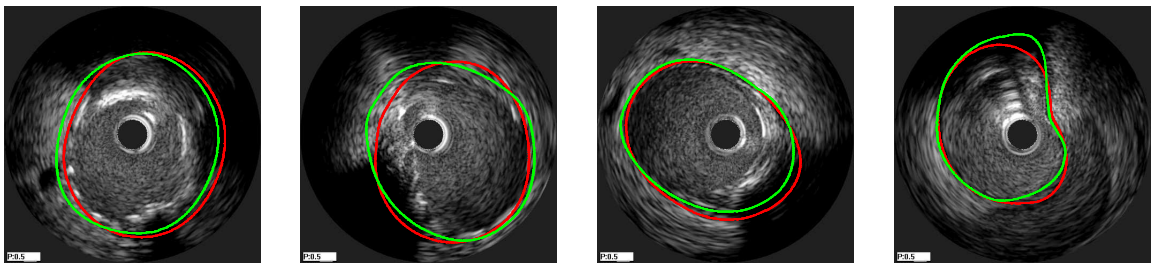


Figure 3: Comparison between groundtruth (green) and the proposed method (red).

[2] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *T-PAMI*, 26(9):1124–1137, 2004.

[3] M. R. Cardinal and et al. Intravascular ultrasound image segmentation: a three-dimensional fast-marching method based on gray level distributions. *T-MI*, 25(1):590–601, 2006.

[4] T. Chan and W. Zhu. Level set based shape prior segmentation. In *CVPR*, 2005.

[5] Esmeraldo Filho and et al. Detection & quantification of calcifications in ivus by automatic thresholding. *Ultrasound in Medicine and Biology*, 34(1):160–165, 2008.

[6] D. Freedman and T. Zhang. Interactive graph cut based segmentation with shape priors. In *CVPR*, pages 755–762, 2005.

[7] J. D. Klingensmith, R. Shekhar, and D. G. Vince. Evaluation of three-dimensional segmentation algorithms for the identification of luminal and medial adventitial borders in intravascular ultrasound images. *T-MI*, 19(10):996–1011, 2000.

[8] Kang Li, Xiaodong Wu, Danny Z. Chen, and Milan Sonka. Optimal surface segmentation in volumetric images-a graph theoretic approach. *T-PAMI*, 28(1):119–134, 2006.

[9] J. Malcolm, Y. Rathi, and A. Tannenbaum. Graph cut segmentation with nonlinear shape priors. In *ICIP*, pages 365–368, 2007.

[10] M. Mulet-Parada and J. Noble. 2D + T acoustic boundary detection in echocardiography. *MIA*, 4 (1):21–30, 2000.

[11] Qi Song, Xiaodong Wu, Yunlong Liu, Mona Garvin, and Milan Sonk. Simultaneous searching of globally optimal interacting surfaces with shape priors. In *CVPR*, pages 2879–2886, 2010.

[12] Milan Sonka and et al. Segmentation of intravascular ultrasound images: A knowledge-based approach. *T-MI*, 14:719–732, 1995.

[13] Gozde Unal and et al. Shape-driven segmentation of the arterial wall in intravascular ultrasound images. *IEEE Trans. Info. Tech. in Biomed.*, 12(3):335–347, 2008.

Motion Analysis

3D Motion Estimation of Carpal Bones from Single View Fluoroscopic Sequences

Xin Chen[1]
xin.chen@manchester.ac.uk
Jim Graham[1]
jim.graham@manchester.ac.uk
Charles Hutchinson[2]
c.e.hutchinson@warwick.ac.uk
Lindsay Muir[3]
lindsay.muir@srft.nhs.uk

[1] Imaging Sciences Research Group
University of Manchester
Manchester, UK

[2] Clinical Sciences Research Institute
University of Warwick
Coventry, UK

[3] Consultant Orthopaedic Surgeon,
Salford Royal Hospital
Salford, UK

Abstract

We present a novel framework for inferring 3D carpal bone kinematics and bone shapes from a single view fluoroscopic sequence. A hybrid statistical model representing both the kinematics and shape variation of the carpal bones is built, based on a number of 3D CT data sets obtained from different subjects at different poses. Given a fluoroscopic sequence, the wrist pose, carpal bone kinematics and bone shapes are estimated iteratively by matching the statistical model with the 2D images. A specially designed cost function enables smoothed parameter estimation across frames. We have evaluated the proposed method on both simulated data and real fluoroscopic sequences. It was found that the relative positions between carpal bones can be accurately estimated, which is potentially useful for detection of conditions such as scapholunate dissociation.

1 Introduction

The diagnosis of wrist pain is frequently achieved by inspection of 2D video fluoroscopic sequences showing movement of the hand in radial-ulnar or flexion-extension motion. This qualitative interpretation requires the inference of the 3D movement of the carpal bones from the 2D sequences, and requires considerable experience from the practitioner. We seek to achieve a quantitative analysis by computer interpretation. During wrist movement, the eight carpal bones follow a complex, multi-dimensional trajectory, making interpretation of radiographs difficult. For this study we have trained a hybrid statistical model (SM) from a set of CT images from different subjects at different poses. Subsequently, the full 3D carpal bone motions can be recovered by matching the SM with the fluoroscopy sequences through 3D-2D image registration techniques. A number of studies have sought to represent the carpal kinematics using CT or MR data. Van deGiessen et al. [6] presented a 3D rigid registration method based on segmented meshes, which aims to build SM of carpal bones. More recently, they introduced a 4D statistical motion model that locally describes the movement patterns of the carpal bones [7]. 3D-2D registration has been the subject of many studies (e.g. [5]), mainly in the field of registration of pre-operative MR or CT images to intra-operative 2D

images. Our work differs from the above in that we seek to achieve registration of a 2D image sequence to a 3D model (not derived from the same individual) to derive the kinematics of an individual wrist.

The main contributions of this paper, distinguishing it from the aforementioned studies, are: (1) A hybrid SM is developed representing both the complex kinematics and shape variation of the eight carpal bones plus radius and ulna. (2) The full 3D motion and bone shapes are recovered by matching the SM with a single view fluoroscopic sequence: a difficult ill-posed problem. (3) Our initial results show that the relative positions between the carpal bones can be estimated accurately through the proposed framework. We are not aware of any study which attempts to make a 2D to 3D inference in a system of this level of complexity.

The system consists of a training phase and a 3D-2D image registration phase. We currently have CT data from 10 subjects, each at five poses (neutral pose and two extreme poses in flexion-extension and radial-ulnar deviation). In the training phase, only the data from the neutral pose and two extreme poses in the radial-ulnar movement were used, as the radial-ulnar movement is the current concern of this paper. The segmentation of each bone and rigid registration parameters that align bones at different poses within and across the subjects were obtained using an iterative segmentation and registration algorithm [2]. A hybrid statistical model, representing both the kinematics and shape variation, was built from the results of the segmentation-registration framework. The kinematic model was built based on the transformation parameters, while the segmentation result was used to build the statistical shape model for each individual bone. Three sets of parameters need to be estimated during image registration in order to interpret the true 3D motion of the carpal bones: (1) Rigid transformation parameters of the wrist and a global scale factor, denoted by θ=$\{tx, ty, tz, rx, ry, rz, s\}$. tx and tz are the in-plane translations in AP view, and ty is the out-of-plane translation. $r = [rx, ry, rz]^T$ denotes the bone orientations, represented by Rodrigues parameter [3]. The magnitude of vector r is the rotation angle around the axis represented by the normalised unit vector of r. s controls the distance between the centroid of each bone to the origin in the radius, and the global size of the bones. (2) Kinematic model parameter b^m that produces valid poses of the carpal bones during movement. (3) Shape model parameters b_i^q and scale factor s_i for each bone (i). In the 3D-2D image registration phase, the three sets of parameters were estimated in sequence from each frame of the fluoroscopy sequences. Detailed descriptions are given in the following sections.

2 Training of Kinematic Model and Shape Model

We use the six rigid transformation parameters for each bone to train the kinematic model. The common coordinate system for all poses has an origin at a specified point in the radius for one subject. The pose of one subject is described by (tx_1, ty_1, tz_1, rx_1, ry_1, rz_1, ..., tx_{10}, ty_{10}, tz_{10}, rx_{10}, ry_{10}, rz_{10})^t. (8 carpal bones, 1 radius and 1 ulna). The orientation parameter allows for a continuous description of the wrist movement. Then the kinematic model can be parameterised as,

$$M = \mu^m + \phi^m b^m \tag{1}$$

where the mean pose μ^m (m is a notation indicating the model parameters) and the principal subspace matrix ϕ^m are computed from 3 (poses)$\times$ 10 (subjects) training samples using PCA. The vector b^m represents the kinematic parameters that describe the pose of M along each principal direction. In our experiments, the first 8 significant components are used, which keeps 98% of variation.

The statistical shape model of each bone is a point distribution model, built using the segmented volume of the same training subjects. The 3D structure of each bone is described by a set of approximately 1000 points on the segmented surface. Correspondence between these points across subjects was established by the minimum description length algorithm [4]. The deformable shape model is then described as,

$$Q_i = \mu_i^q + \phi_i^q b_i^q \tag{2}$$

where μ_i^q and ϕ_i^q (q is a notation indicating the shape parameters) are the mean shape and the principal subspace matrix for the i^{th} bone. b_i^q is the shape model parameter to be estimated. In order to keep the complexity within limits, only the first 3 significant components are used which keeps 84% of variation.

Based on the point distribution model of each bone and the kinematic model, a hybrid statistical mesh model can be built by using the Crust mesh construction algorithm [1]. Figure 1 shows the poses of the first component of the kinematic model (represented by the mean shapes of each bone) and the shapes of the first component of the scaphoid.

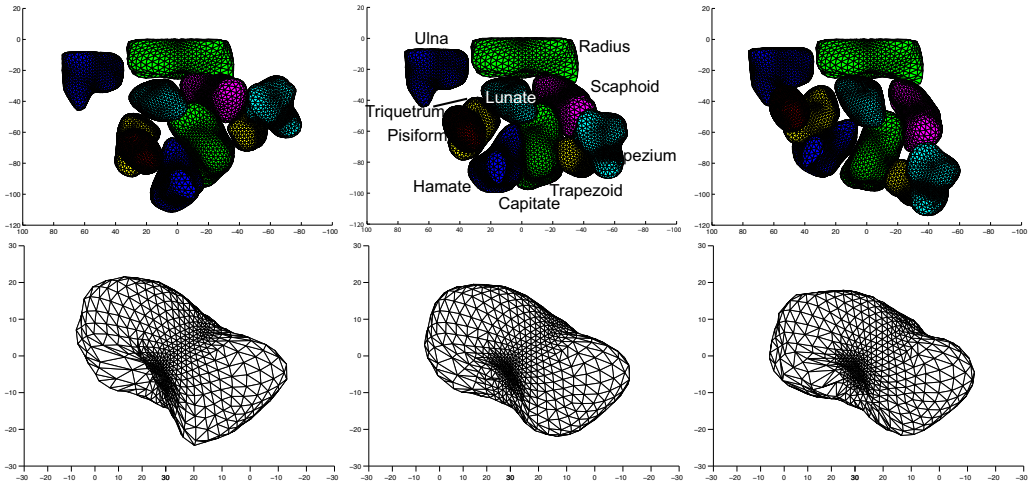


Figure 1: Top row: The poses of the first component of kinematic model. Bottom row: the first component of the shape model of the scaphoid. In each case the mean +/-1.5s.d. are shown.

3 3D-2D Image Registration

The statistical mesh model from the training data is then used to match with the fluoroscopic sequence to infer the 3D motion and bone shapes. Figure 2(a) summarises the registration process. Detailed descriptions are given in the following subsections.

3.1 Cost function

Preprocessing of the fluoroscopic image consists of normalising the local intensity to zero mean and unit standard deviation, followed by anisotropic diffusion. Figure 2(b) shows an example of the gradient magnitude map of the fluoroscopic image after enhancement. To optimise the pose parameters we iteratively generate projections from the mesh model with updated model parameters. Taking the mesh model as a binary volume, the projected intensity is negatively proportional to the sum of binary values along the ray from the source to each pixel in the image plane. Figure 2(b) shows an example.

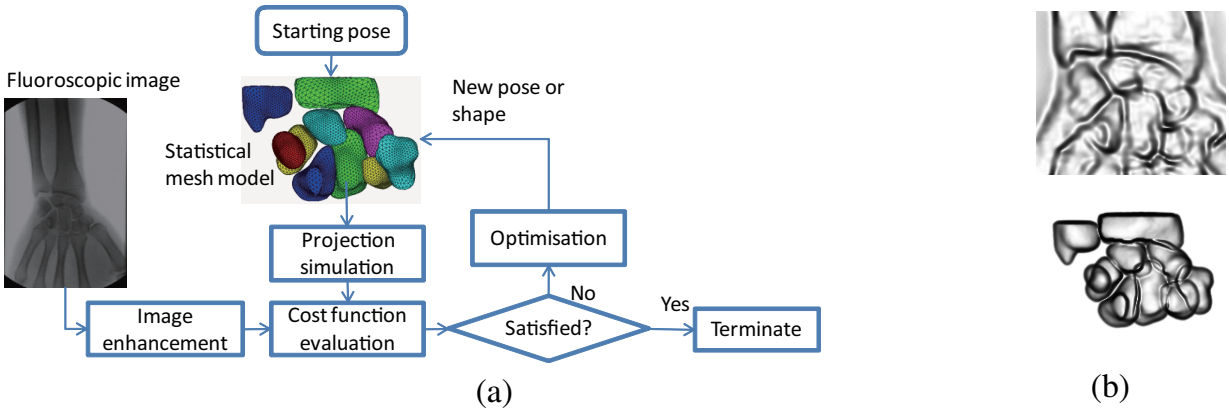


Figure 2: (a) Overview of the 3D-2D image registration process. (b) The gradient magnitude map of the fluoroscopic image after enhancement (cropped to show the region of interest) (top) and the simulated image from mesh model (bottom).

To evaluate the similarity between the fluoroscopic image and the simulated image, we use the cost function shown in Eqn. (3), based on the gradient along horizontal and vertical directions as well as the gradient magnitude of the two images. The adjacent frames of the current fluoroscopic image were also taken into account in the cost function to make the estimated poses smooth across frames.

Taking C(A,B) as the Normalised Correlation Coefficient between two images A and B, we can write the cost function as:

$$E = C(Om_{k-1}, Om_k) + \sum_{p=k-1,k,k+1} w_p(C(Im_p, Dm_k) + C(Ix_p, Dx_k) + C(Iy_p, Dy_k)) \quad (3)$$

where k is the current frame number of the fluoroscopic sequence. Im_p, Ix_p and Iy_p are the gradient magnitude, vertical gradient and horizontal gradient of the fluoroscopic image at the p^{th} frame respectively. Dm_k, Dx_k and Dy_k are the corresponding values of the simulated image. The second term calculates a cross-correlation between sets of three adjacent frames with weights w_{k-1}, w_k and w_{k+1}= 0.2, 0.6, 0.2 respectively, making the estimated pose smooth across frames. For the first term of the cost function, the vertices in the statistical mesh model are projected to the image plane; we assume the intensities at those projected points are similar across adjacent frames. Om_{k-1} and Om_k represent the gradient magnitude of the previous frame and the current frame at the projected correspondence positions. The first term makes the shape of the cost function sharper, leading to a faster and more accurate optimisation result. The $(k-1)^{th}$ frame and $(k+1)^{th}$ frame are not evaluated for the first and last frame respectively.

3.2 Optimisation

The optimisation method used is the best neighbour search combined with parabola fitting. The multi-dimensional search space (θ, b^m and b^q) is explored iteratively by individual 1D line search. The cost function is evaluated at the current position, positive and negative neighbour positions (defined by a search range), then an optimum is found by fitting a parabola to the 3 evaluated positions, and iteratively refined by reducing the search range until convergence.

In our case, the true sizes of the bones are unknown; recovering the 3D pose from a single image is therefore a difficult, ill posed, problem. Any movement along the out-of-plane translation, could be compensated by scaling of the bone. We estimate parameters (except ty) in the following sequence: tx, tz, rx, ry, rz, b^m, s, s_i and b_i^q. After convergence, the

estimated pose of the current frame is used as the starting pose for the next frame. The shape model parameters b_i^q are only estimated once in the first frame, as these are not improved significantly when we include more frames and the fitting is made significantly more complex and time consuming. The whole process was performed in a 3-level multi-scale framework at each frame to enhance the robustness of the registration.

4 Evaluation

As there is no ground truth relating the fluoroscopic sequences to the CT data, we evaluated our framework based on a number of simulated fluoroscopic sequences generated from the 3D CT data. All CT volumes have been resampled to an isocubic volume with voxel dimension of 0.5 mm. We linearly interpolated a number of poses between the neutral pose and two extreme poses of radial-ulnar deviation in a full movement cycle containing 50 poses for each of 10 subjects. The ray-casting method was then used to generate a simulated fluoroscopic sequence from those interpolated 3D data. We evaluated the proposed framework in the leave-one-out manner. The 3D pose of the simulated test subject was then calculated as described in section 3, and registration error measured by the 3D Euclidian distance of each corresponding point of the mesh between the target pose and the estimated pose is presented in Table 1. The error of the registration is mainly caused by the ill posed problem (confusion between the scale and translation along the out-of-plane direction), whereas the errors along the in-plane directions are very small with average error of about 1 mm and maximum error within 2mm.

It is important to mention that the relative positions of the carpal bones with respect to each other can be estimated much more accurately than the absolute positions of the individual bones. The registration error of the 3D distance between the centroid of Triquetrum and the centroid of Lunate (dTL), and the distance between the centroid of Lunate and the centroid of Scaphoid (dLS) were also measured. The errors are 0.59±0.37 mm and 0.91±0.50 mm for dTL and dLS respectively, compared to a bone size of about 15-20 mm. One of the conditions that may be assessed using this method is dissociation, where the distance between the bones is larger than normal. Scapholunate dissociation is one of the most common of these. We normalise the dLS by dividing it by the estimated global scale factor s and an average of the local scale factors for lunate and scaphoid. From the tested 10 subjects, we successfully identified the subjects suffering from scapholunate dissociation (dLS=19.39±0.77 mm) from the normal subjects (dLS=17.25±0.42 mm). Making this type of measurement without a 3D statistical model would be impossible.

Table 1: The average error in mm, measured in 3D, between the target and estimated correspondence points of each carpal bone of 10 subjects: Triquetrum(Tri), Lunate(Lun), Scaphoid(Sca), Pisiform(Pis), Hamate(Ham), Capitate (Cap), Trapezoid (Trd) Trapezium (Trm). The measurement errors of dTL and dLS.

	eTri	eLun	eSca	ePis	eHam	eCap	eTrd	eTrm	Total	eTL	eLS
Err3D	2.7±1.3	2.6±1.3	3.3±1.8	3.4±1.9	3.3±1.9	3.3±2.0	3.3±2.3	3.8±2.2	3.2 ±1.9	0.59±0.37	0.91±0.50
ErrX	0.8±0.7	1.0±0.8	1.1±0.9	1.2±1.0	0.9±0.7	1.1±0.8	0.9±0.7	1.1±0.9	1.1 ± 0.9	/	/
ErrY	1.9±1.4	1.5±1.3	2.5±1.9	2.4±2.0	2.7±2.1	2.7±2.2	2.8±2.5	3.1±2.3	2.3±2.0	/	/
ErrZ	1.3±1.0	1.3±1.0	1.1±0.9	1.3±1.1	1.3±0.7	0.9±0.7	0.8±0.6	1.1±1.0	1.2±1.0	/	/

We also tested our framework on real fluoroscopic sequences. Although the matching error cannot be quantified, the registration results show good visual correspondence and have been confirmed by a clinician. A sample frame of the matching result and the corresponding 3D pose are shown in Fig. 3 in which the projected contours from the 3D mesh model are superimposed on the preprocessed fluoroscopy image. The estimated 3D mesh model in the

palmar and dorsal views are shown in middle and right respectively.

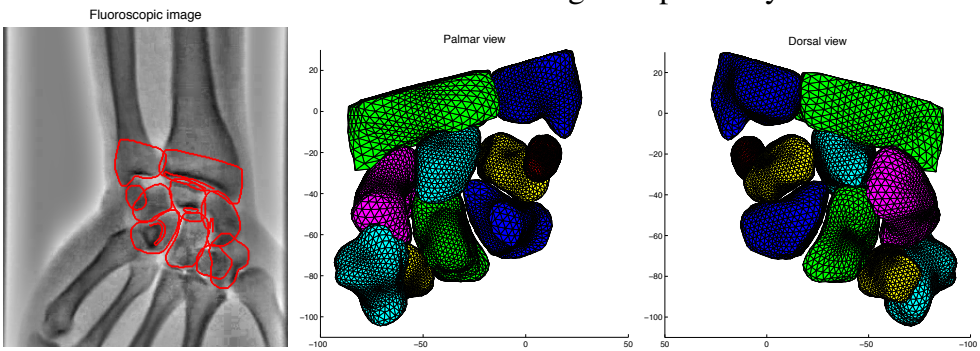


Figure 3: Registration result of one frame from a real fluoroscopic sequence.

5 Concluding Remarks

We have presented a complete framework that is able to infer the 3D motion of carpal bones from a single view fluoroscopic sequence. It uses a hybrid statistical model to estimate both the kinematics and bone shapes from the fluoroscopic sequences allowing the motion of carpal bones during radial-ulnar deviation to be estimated. Particularly, the relative positions between carpal bones can be estimated accurately. This is potentially useful for detection of dissociation conditions, such as scapholunate dissociation, where the underlying pathology is a rupture of one or more ligaments, and the diagnosis rests on a judgement regarding the joint separation.

In further work we will extend the current statistical model with more training data (in progress) and test the framework for the flexion-extension movement.

References

[1] N. Amenta. The crust algorithm for 3D surface reconstruction. In *Proc. syposium on computational geometry*, 1999.

[2] X. Chen, J. Graham, and C.E. Hutchinson. Integrated framework for simultaneous segmentation and registration of carpal bones. In *Proc. ICIP*, pages 433–436, 2011.

[3] J. J. Craig. *Introduction to Robotics: Mechanics and Control*. Addison-Wesley Publishing, 1989.

[4] R.H. Davies, C. Twining, T.F. Cootes, and C.J. Taylor. Building 3-D statistical shape models by direct optimisation. *IEEE T-MI*, 29(4):961–980, 2010.

[5] G. P. Penney, P. G. Batchelor, D. L. G. Hill, D. J. Hawkes, and J. Weese. Validation of a two- to three-dimensional registration algorithm for aligning preoperative CT images and intraoperative fluoroscopy images. *Medical Physics*, 28:1024–1032, 2001.

[6] M. van deGiessen, G. J. Streekstra, S. D. Strackee, M. Maas, K. A. Grimbergen, L. J. van Vliet, and F. M. Vos. Constrained registration of the wrist joint. *IEEE T-MI*, 28(12): 1861–1869, 2009.

[7] M. van deGiessen, M. Fournani, F. M. Vos, S. D. Strackee, M. Maas, L. J. van Vliet, K. Grimbergen, and G. J. Streekstra. A 4D statistical model of wrist bone motion patterns. *IEEE T-MI, online preprint, october*, 2011.

Tracking neutrophils in zebrafish: the use of synthetic data sets

Constantino Carlos Reyes-Aldasoro[1]
c.c.reyes-aldasoro@sussex.ac.uk
Katy Henry [2]
k.henry@sheffield.ac.uk
Stephen A. Renshaw[2]
s.a.renshaw@sheffield.ac.uk

[1] Biomedical Engineering School of Engineering and Informatics University of Sussex, UK

[2] MRC Centre for Developmental and Biomedical Genetics, and Department of Infection and Immunity, University of Sheffield, UK

Abstract

In this work we present a series of data sets that model the behaviour of neutrophils as observed with a confocal microscope. The data sets describe important characteristics of the migration of neutrophils such as collisions and path tortuosity as well as different levels of background noise. Neutrophil trajectories were manually defined, and Gaussian shapes similar to those of real data sets were assigned to each position of a neutrophil. The availability of synthetic data sets such as the ones proposed here, together with appropriate *gold standards* will benefit those wanting to test the robustness and accuracy of segmentation and tracking algorithms.

1 Introduction

The advent of multiphoton and confocal fluorescence microscopy, which allows 3D imaging of specimens *in vivo* with high spatial and temporal resolution, has been widely adopted in the Life Sciences [6] for applications ranging from the observation of microvascular permeability [11], the assessment of mitochondrial function [15], and examining tumour angiogenesis [13], to the observation of neutrophil apoptosis and migration [3]. Confocal and multiphoton microscopes capture the intensity value (i) at a specific three dimensional location (x,y,z). This intensity is related to the photons detected at the camera, which are in turn related to the structure or concentration of a substance within the sample which can present different fluorescent reactions to the excitation frequency (f). As the observation is repeated over time (t), the data becomes a 5-dimensional matrix $i(x,y,z,f,t)$ (Fig. 1a). Thus, the data acquired on a single experiment can easily be of many gigabytes (GB) and special considerations have to be implemented for the storage and transport of such data sets.

Zebrafish larvae have emerged as a key model organism in recent years, with a unique combination of advantages over other model systems for the detailed study of inflammation biology *in vivo* [9]. Their optical transparency allows visualisation of physiological and pathological processes as it is possible to visualise both the anatomy of the fish together with fluorescent cell populations *in vivo* (Fig. 1b). Genetic manipulations can be easily performed, both to genetically manipulate the inflammatory response, but also to label individual cells

with fluorescent markers [10]. This then allows these cell populations to be observed in high temporal and spatial resolutions, during inflammation resolution, using multiphoton and confocal microscopy.

The combination of confocal microscopes with zebrafish models has provided a crucial setting for the detailed study of inflammation biology [2]. Inflammation is a process critical to life itself, without it multicellular animals could not protect themselves against the threat of competing unicellular microorganisms. However, the power of our immune systems can also be a threat to our own health when inflammatory responses fail to resolve in a timely fashion. Understanding how immune cells behave during all phases of inflammation *in vivo* is an important part of understanding cellular migration and interaction. Neutrophils and macrophages (collectively termed *myeloid cells*) are the key cells of the innate immune system, and are essential aspects of both host defence against infectious disease, and tissue damage caused during inflammatory disease. The function of such cells is notoriously difficult to examine *in vivo*.

Neutrophil analysis from multiphoton data sets is complicated for several reasons. First, the data sets can be of high-resolution images as mentioned previously and thus the computational complexity is considerable. Second, neutrophil behaviour is typically analysed from derived measurements, *e.g.* velocity, shape changes or activation, rather than from direct observation. Thus it is necessary to pre-process the data: namely, to segment the neutrophils and track them in time [12] before any quantitative analysis can be performed. Third, the collisions that occur between neutrophils present a major challenge to tracking, as two separate fluorescent cells may appear as a single large fluorescent cell. The problem is further complicated by the fact that neutrophils can elongate considerably, and two pseudopods can stretch apart looking then as two separate neutrophils since the region between the two elongated pseudopods will decrease in fluorescence intensity. Fourth, *gold-standards* to evaluate the robustness of segmentation and tracking algorithms are rarely available for biomedical data, as these require considerable time from an expert user to manually annotated the data, in this case segment and track neutrophils. To the best of the authors' knowledge, there are no public data sets of neutrophils with gold-standards that can be used to compare and validate analysis algorithms. Furthermore, many labs employ proprietary software with its own routines, which makes replication and verification difficult [7] as these routines are not available to the wider community.

For this work, we generated a series of synthetic data sets in MATLAB © that simulate a group of migrating neutrophils. The neutrophils exhibit different migration behaviours, such as change in shape, collisions in their paths and different levels of tortuosity. The data sets and their trajectories are freely available [14] and can be used as a benchmark to test the robustness of segmentation and tracking algorithms for neutrophils in zebrafish.

2 Synthetic Neutrophil Data Sets

Numerous data sets of neutrophils in zebrafish were carefully observed to capture the main characteristics to be reproduced in the artificial sets. Six trajectories were manually determined by setting the $row, column$ positions of the centroids at every time point for 98 time frames with z-stacks of 11 slices and 550×550 pixels. Each trajectory was designed so that it would represent different neutrophil behaviours: some trajectories were very oriented and had movements with uniform distance between time frames, whilst others were less uniform and would move at different velocities, some were tortuous whilst others were straight. The

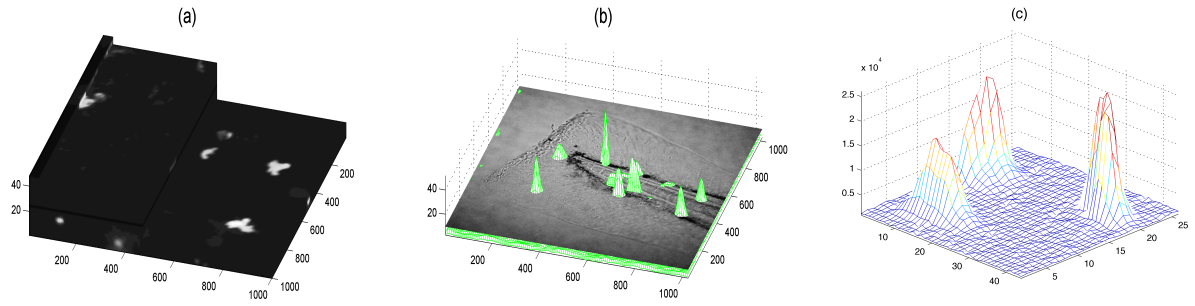


Figure 1: Fluorescent neutrophils in zebrafish: (a) One time frame t, 24 slices (1024 $\times$ 1024 pixels) in the z axis, two frequencies (f_1 in the top for neutrophils, f_2 in the bottom for macrophages). The data has been sliced to visualise the intensity at different levels. (b) A single fluorescent slice plotted as a mesh with intensity-related height, and one slice of the DIC image, which shows the anatomy of the fish. (c) A region of interest showing the decaying intensity of the neutrophils, which will be modelled as a Gaussian function.

trajectories of cells 1 and 2 collided several times in the second half of the time frames whilst cells 3 and 4 collided at the beginning of the movement. Cell 6 migrated without meandering and then stopped at the end (which represents the wound area of an inflammation-based experiment) whilst 5 presented a delayed activation. The neutrophils were then assigned intensities that modelled the fluorescent intensity of the neutrophils with a 3D Gaussian distribution. The orientation of the Gaussians varied according to the displacement of the artificial neutrophils, *i.e.* they were round when the cells were static, or elongated when in movement. The shapes were simplistic and in the future we plan to assign more realistic shapes with extending pseudopods.

The tracks with the Gaussians were saved as the *gold standard* and five different data sets were generated by adding varying levels of white Gaussian noise resulting in data sets with distributions with increasing similarity between the neutrophils and the background reflected by the decreasing values of the Bhattacharyya Distance (1.61, 1.25, 1, 0.66, 0.45) (Fig. 2b). The Bhattacharyya Distance (*BD*) [1, 5] is a well-known measurement of the separability between two probabilistic distributions based on their means and standard deviations.

In its simplest formulation, the Bhattacharyya distance between two classes can be calculated from the variance and mean of each class in the following way:

$$D_B(k_1,k_2) = \frac{1}{4}\ln\left(\frac{1}{4}\left(\frac{\sigma_{k_1}^2}{\sigma_{k_2}^2}+\frac{\sigma_{k_2}^2}{\sigma_{k_1}^2}+2\right)\right)+\frac{1}{4}\left(\frac{(\mu_{k_1}-\mu_{k_2})^2}{\sigma_{k_1}^2+\sigma_{k_2}^2}\right) \tag{1}$$

where: $D_B(k_1,k_2)$ is the Bhattacharyya distance between k_1-th and k_2-th classes, σ_{k_1} is the variance of the k_1-th class, μ_{k_1} is the mean of the k_1-th class, and k_1,k_2 are two different training classes. For the multidimensional distance, the variances are replaced by co-variance matrices and the means become vectors [4]:

$$D_B(k_1,k_2) = \frac{1}{2}\ln\left(\frac{|\frac{1}{2}(\Sigma_{k_1}+\Sigma_{k_2})|}{\sqrt{|\Sigma_{k_2}||\Sigma_{k_1}|}}\right)+\frac{1}{4}(\mu_{k_1}-\mu_{k_2})^T(\Sigma_{k_1}+\Sigma_{k_2})^{-1}(\mu_{k_1}-\mu_{k_2}) \tag{2}$$

where $(\cdot)^T$ is the transpose of the matrix. The Mahalanobis distance used in Fisher Linear Discriminant Analysis (LDA) is a particular case of the Bhattacharyya, when the variances of the two classes are equal; this would eliminate the first term of the distance.

Fig. 2c,d display one slice of the z-stack for time frame 26 for the data sets with largest *BD* (less noise) and smallest *BD* (more noise). The mean intensity values for the background and neutrophil classes were 10 and 18 respectively.

3 Segmentation and Tracking Algorithm

To test the data sets, we implemented a hysteresis segmentation and keyhole model of movement tracking algorithm as proposed in [12]. Briefly, the segmentation algorithm performed intensity-based segmentation with a hysteresis double threshold, with high and low thresholds determined based on Otsu's algorithm [8]. Volume elements (voxels) below and above the lower threshold were classified as background and neutrophils respectively. Intermediate-level voxels were classified as cells if in contact with voxels above the high threshold, or as background otherwise. Tracking was performed with a keyhole model of movement where cell movement is predicted to follow one of two possibilities: either completely random around the current position of the cell, or oriented following the same orientation and velocity as the previous jump. These options are translated into two regions of probability: a circle and a wedge, together resembling a keyhole. Several post-processing tasks were performed: initial nodes were validated with a backward linking of nodes, tracks for which neutrophils disappear for one frame were joined, small tracks were discarded, and collisions between neutrophils were detected and cells split to be re-tracked.

4 Results

The segmentation and tracking algorithms were applied to the 5 data sets described above and a sensitivity analysis for the threshold levels was performed, as this was the only input parameter that can be modified by the user. We defined two measurements of accuracy: (1) the number of tracks generated by the algorithms (Fig. 3a) and (2) the distance between the centroids of the segmented neutrophils and the *gold standard* (Fig. 3b). It was important to use both measurements as it is possible to have scenarios where tracks can be broken into several segments, each of which will be very close to the gold standard (good performance in 2, bad performance in 1) or conversely the tracking could confuse neutrophils after a collision and follow the incorrect neutrophil along the track of the gold standard (good performance in 1, bad performance in 2). The threshold values automatically calculated were $[19.8, 21.1]$, therefore, we varied the levels from 11 to 31 for the lower threshold and 13 to 33 for the higher threshold. As it should be expected, the errors increase at the extreme values: background regions were considered neutrophils with low threshold levels, and faint neutrophils were lost with high threshold levels. However, the results are fairly stable for a good range of threshold values ($[19-27]$ for the lower threshold, shaded areas) where the number of tracks is $(6-8)$ for all the data sets and the average distance between the centroids of the segmented neutrophils and the gold standard is 1.09 pixels.

5 Conclusions

In this work we present synthetic data sets that simulate the migration of neutrophils as observed with a confocal microscope. We tested the data sets with a segmentation and tracking algorithm which provided stable results for a large range of threshold values. In the future,

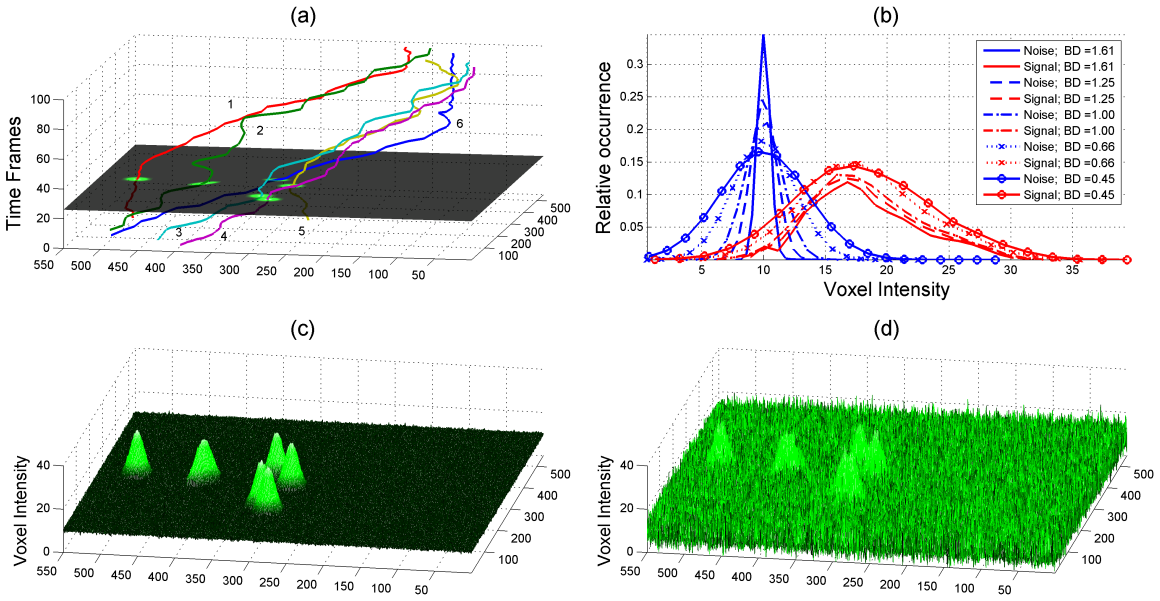


Figure 2: Synthetic data sets. (a) One slice of the intensity at time 30 and the paths of six individual neutrophils shown as coloured lines. The z-axis indicates time. (b) Five sets of histograms of the distributions for background (noise) and neutrophils (signal) for different levels of noise. The separability of the two classes is indicated by the values of the Bhattacharyya Distance (*BD*). (c) One slice of the set with $BD = 1.61$ shown as a mesh where the intensity corresponds to the z-axis. (d) One slice of the set with $BD = 0.45$. The level of the noise can be easily compared between the two data sets.

we plan to develop more data sets that describe different neutrophil behaviour characteristics as well as acquisition settings, *i.e.* different resolutions, as validation data sets to test the robustness of analysis algorithms. We also plan to create and release a function through which new tracks with different behaviours can be generated as required. The data sets and the neutrophil trajectories are freely available [14] .

References

[1] Guy B Coleman and Harry C Andrews. Image Segmentation by Clustering. *Proceedings of the IEEE*, 67(5):773–785, 1979.

[2] P. M. Elks, C. A. Loynes, and S. A. Renshaw. Measuring inflammatory cell migration in the zebrafish. *Methods Mol Biol*, 769:261–75, 2011.

[3] P. M. Elks, F. J. van Eeden, G. Dixon, X. Wang, C. C. Reyes-Aldasoro, P. W. Ingham, M. K. Whyte, S. R. Walmsley, and S. A. Renshaw. Activation of hif-1alpha delays inflammation resolution by promoting neutrophil persistence in a zebrafish inflammation model. *Blood*, 118:712–722, 2011.

[4] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press, 1972.

[5] Thomas Kailath. The divergence and bhattacharyya distance measures in signal selection. *IEEE Trans. on Commun. Technology*, 15(1):52–60, 1967.

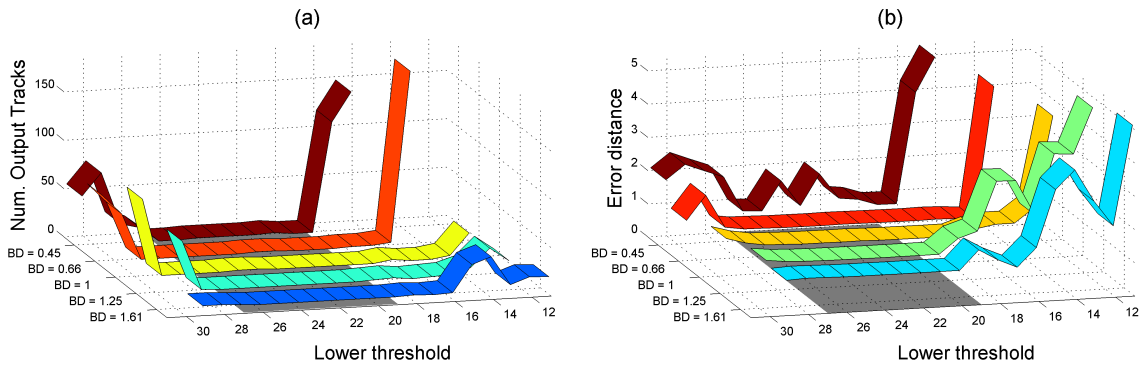


Figure 3: Validation of the algorithms against different input thresholds. (a) Number of tracks generated by the algorithms; the correct number was 6. (b) Average distance between the centroids of the segmented neutrophils and the gold standard in pixels. The shaded are denotes the range in which the algorithms provide consistent results.

[6] K. Konig. Multiphoton microscopy in life sciences. *J Microsc*, 200(Pt 2):83–104., 2000.

[7] T. A. Nielsen, H. Nilsson, and T. Matheson. A formal mathematical framework for physiological observations, experiments and analyses. *J R Soc Interface*, 2011.

[8] N. Otsu. A threshold selection method from gray level histograms. *IEEE Trans. Syst. Man Cybern.*, 9:62–66, 1979.

[9] S. A. Renshaw and N. S. Trede. A model 450 million years in the making: zebrafish and vertebrate immunity. *Disease Models & Mechanisms*, 5(1):38, 2012.

[10] S. A. Renshaw, C. A. Loynes, D. M. Trushell, S. Elworthy, P. W. Ingham, and M. K. Whyte. A transgenic zebrafish model of neutrophilic inflammation. *Blood*, 108(13): 3976–8, 2006.

[11] C. C. Reyes-Aldasoro, I. Wilson, V. E. Prise, P. R. Barber, S. M. Ameer-Beg, B. Vojnovic, V. J. Cunningham, and G. M. Tozer. Estimation of apparent tumor vascular permeability from multiphoton fluorescence microscopic images of p22 rat sarcomas in vivo. *Microcirculation*, 15(1):65–79, 2008.

[12] C.C. Reyes-Aldasoro, Y. Zhao, D. Coca, S.A. Billings, V. Kadirkamanathan, G. M. Tozer, and S. A. Renshaw. Analysis of immune cell function using in vivo cell shape analysis and tracking. In *4th IAPR Int Conf Pat Rec Bioinformatics*, Shef, UK, 2009.

[13] G. M. Tozer et al. Blood vessel maturation and response to vascular-disrupting therapy in single vascular endothelial growth factor-A isoform-producing tumors. *Cancer Res.*, 68(7):2301–2311, 2008.

[14] Webpage. http://www.neutrophil.org.uk. 2012.

[15] Z. Zhong et al. Nim811, a mitochondrial permeability transition inhibitor, prevents mitochondrial depolarization in small-for-size rat liver grafts. *American Journal of Transplantation*, 7(5):1103–1111, 2007.

Poster Session 1

Skin Lesion Image Recognition with Computer Vision and Human in the Loop

Orod Razeghi[1]
http://www.cs.nott.ac.uk/~ozr
Guoping Qiu[1]
http://www.cs.nott.ac.uk/~qiu
Hywel Williams[2]
Hywel.Williams@nottingham.ac.uk
Kim Thomas[2]
Kim.Thomas@nottingham.ac.uk

[1] VIPLAB, IMA Group
Computer Science
University of Nottingham
Nottingham, UK

[2] Centre of
Evidence-Based Dermatology
University of Nottingham
Nottingham, UK

Abstract

In this paper, we present an interactive human in the loop computer vision technique for the recognition of skin lesion images. We have designed a dermatology "Question and Answer" bank suitable for interactively extracting human perceptual knowledge of images in order to assist computer vision algorithms in boosting recognition accuracies. We present experimental results to show that for some diseases, traditional computer vision techniques can only achieve a recognition rate of 20%, whilst with human in the loop the performance can be boosted to over 96%. We also show that users do not require any medical knowledge to answer these questions to achieve excellent recognition rates.

1 Introduction

Recently, researchers have advocated and developed the so-called interactive imaging and vision [11] or human in the loop [4] approach to visual object recognition. It has been well known that some problems, that are difficult for computer to solve, are actually very simple for human. For instance, fully automated and accurate computer visual object recognition has proved to be very difficult, if not entirely impossible. An intermediate solution to develop useful and practical technology is to make human and computer work in harmony and exploit their respective strengths. We believe that the approach introduced in [4] is particularly suitable for medical image recognition applications, such as computer aided diagnosis.

Unlike the majority of publications in the area of computer vision for dermatology applications, this paper is the first that attempts to apply promising human in the loop visual recognition technique to automatic recognition of various conditions, including non-melanoma skin diseases, and is certainly one that achieves the best results in the literature. Our contributions include: i) demonstrated for the first time that human in the loop visual recognition can significantly boost accuracy and achieve near perfect recognition results, ii) designed a dermatology relevant "Q&A" bank containing 21 questions and over 100 possible answers suitable for human in the loop visual recognition solutions and iii) shown that such system can be exploited by users without any medical knowledge.

2 Computer Vision in Dermatological Applications

An example of related work is an image analysis system presented in [1] that differentiates early melanoma from benign pigmented lesions. The analysis system extracts features related to the size, shape, boundary, and colour of each lesion. Another solution in form of an automated melanoma recognition system is presented in [6]. A binary mask of lesion is obtained by a number of basic segmentation algorithms alongside a fusion strategy. A set of shape and radiometric features are calculated to determine the malignancy of a lesion.

A model of tissue colouration is presented in [5]. The model is built by computing the spectral composition of light remitted from normal human skin colour, and comparing it to abnormal tissues. As an alternative the framework in [12] assesses a series of 588 flat pigmented skin lesions. The proposed analyser employs an artificial neural network. A feature selection procedure confirms that as few as 13 variables are adequate to discriminate the two groups of "melanoma" and "other pigmented" skin lesions.

Although the literature demonstrates a number of attempts at fabricating Content Based Image Retrieval (CBIR) Medical Systems for dermatological purposes [3][7], and quite a few attempts at assessing severity of specific skin diseases automatically [13], the lack of a reliable system for unskilled users, or an assistant tool for general practitioners is apparent.

3 Human in the Loop Skin Lesion Recognition

Our system adopts the framework of [4] for incorporating any multi-class object recognition algorithm that produces a probabilistic output over classes, as follows:

$$p(c|x,U) = \frac{p(U|c,x)p(c|x)}{\sum_c p(U|c)p(c|x)} = \frac{p(U|c)p(c|x)}{\sum_c p(U|c)p(c|x)} \tag{1}$$

where c is class, x is image, and U is any random sequence of user answers. The assumption that $p(U|c,x) = p(U|c)$ suggests that the types of randomness present in user answers is class-dependent and not image-dependent.

In our implementation of the above framework, we employed 10 image features including Coloured Pattern Appearance Model (CPAM) [10], Geometric Blur (GB) [15], Global Image Descriptor (GIST) [8], Pyramid Histogram of Oriented Gradients (PHOG) and its variations [15], Scale-invariant Feature Transform (SIFT) and its variations, Pyramid Histogram of Visual Words (PHOW) and its variations [14], and Self-similarity Feature (SSIM) [15]. We used *OBSCURE* [9], a state of the art, publicly available open source Support Vector Machine (SVM) based classifier for the visual classification, $p(c|x)$.

Similar to the original work [4], we also used a multinomial distribution with a Dirichlet prior to model user responses $p(U|c)$, and used KL divergence and maximum information gain to choose the next suitable question. Figure 1 illustrates the entire process.

4 Experimental Results

4.1 Datasets

We have collected two datasets from various Internet sources. The first and second datasets contain 90 and 706 dermatological images from 3 and 7 different skin diseases respectively. The lesions were manually segmented using a bounding box that includes pixels of lesion,

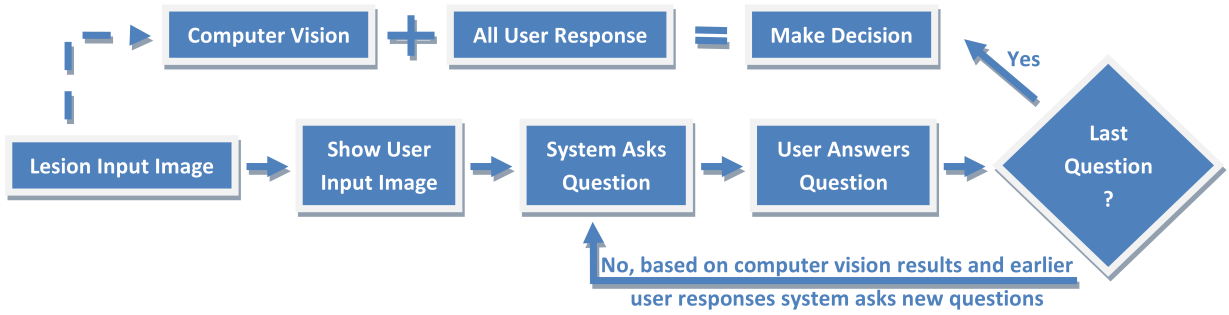


Figure 1: A skin lesion image is displayed to the user. For each image, a question and its possible answers are also displayed. The user answers the question. The user will repeat the process until all the questions are finished or a threshold is reached. The next question is selected by looking at previous user's answers and computer vision input. The final decision is also made by combining these two elements.

healthy skin and noise, such as hair. Features were extracted from the entire bounding box, which as a whole is treated as a single instance. Images with their ground truth classification were mainly collected from http://www.dermis.net.

4.2 Dermatology Question and Answer Bank

A set of questions, which both summaries patient general conditions as well as her skin lesion characteristics, were designed to help with obtaining user perception of patient in a (fabricated) scenario. The questions range from the age of patient to colour and surface features of skin lesions. We have consulted medical professionals and a dermatological reference [2] to scientifically derive these questions.

Table 1 lists 8 questions and 36 possible answers used for testing the first dataset and Table 2 lists 13 questions and 67 possible answers used for testing the second dataset. Wherever specific medical terms are used, a guide image with explanations is available for users to avoid confusion.

4.3 Results

Table 3 shows the results for the 1st and 2nd datasets. In the 1st dataset, 15 randomly selected images from each of the 3 diseases were used in training and the rest were used for testing.

Table 1: Dermatology First Dataset Questions

Question	Attributes (Possible Answers)
01 Site	Head, Trunk, Arms, Legs
02 Condition	Acute, Chronic
03 Surface	Normal, Scaly, Hyperkeratotic, Warty, Crust, Exudate, Excoriated
04 Lesion	Flat, Raised, Fluid Filled, Surface Broken
04 Colour	Pink, Red, Purple, Mauve, Brown, Black, Blue, White, Yellow, ...
06 Age	Infant, Young, Adult, Old
07 Contagiousness	Contagious, Non-contagious
08 Itchiness	Itchy, Non-itchy

Table 2: Dermatology Second Dataset Questions

Question	Attributes (Possible Answers)
01 Age	Infant, Child, Adult, Elderly
02 History	Personal, Family
03 Site	Face, Scalp, Ears, (Mouth, Tongue, Lips), Trunk, Hands, ...
04 Number	Single, Multiple
05 Distribution	Symmetrical, Asymmetrical, Unilateral, Localised, Generalised
06 Arrangement	Discrete, Coalescing, Disseminated, Annular, Linear, Grouped
07 Erythema	Erythematous, Non-erythematous
08 Duration	Acute, Chronic
09 Type	Flat, Raised Solid, Fluid Filled, Cyst, Comedone, Broken Surface
10 Surface	Normal, Abnormal Keratinisation, Scale, Broken, Crust, Shiny, ...
11 Colour	Due to blood (Red, Pink), Due to pigment (Black, Blue), ...
12 Border	Well defined, Poorly defined, Accentuated edge
13 Shape	Round, Irregular, Rectangular, Serpiginous, Dome shaped, ...

The experiment was repeated 5 times and the results in the table is the average over 5 rounds of experiments by a group of non-expert users. Here, it is clear that computer vision performs very badly on the Scabies images, achieving only 33% correct recognition rate. With human in the loop, the correct recognition rate boosts to 93% - a very significant improvement. Average correct recognition rate across the diseases is just over 57% for computer vision only solution but it is boosted to over 97% with human in the loop.

In the 2nd dataset, 30 images from each disease were randomly selected for training and the remaining 496 images were used for testing. Here the computer vision technique can only achieve 20% recognition rate for Mycosis Fungoides, but with human in the loop, the recognition rate is boosted to over 96% - again a dramatic improvement. The average across the diseases for the entire dataset is 61% for computer vision only and 96% for introducing human in the loop.

These results clearly demonstrate the effectiveness of human in the loop technique for recognising skin lesions. Compared to computer vision only solutions, adding human in the loop can dramatically improve the correct recognition rates. Moreover, computer aided diagnosis is known to be capable of reducing subjectivity, thus can reduce inter observer discrepancies, and our highly accurate results are consistent with this conclusion.

It may be argued that contribution of computer vision is unclear, since human responses are adequate to correctly classify a disease. However, as shown in figure 2, computer vision plays an important role in reducing human labour in terms of number of questions necessary to answer, and the time spent on each image to correctly classify a disease. Furthermore, some images cannot be classified correctly without computer vision, even after asking all the questions. There are over 1000 skin conditions worldwide, a fully functioning system will need a question bank of hundreds, if not thousands of questions. Computer vision will be essential in reducing labour and improving diagnosis.

More interestingly, it was also found that users do not have to answer all the questions correctly in order to achieve correct recognition. It was seen that although users' questions are asked in different orders and users' answers to the same questions are different, the algorithm still recognises images successfully. Figure 3 illustrates the frequency of possible answers selected by users in the 2nd dataset.

Table 3: 1st and 2nd Dataset Accuracy Results

Class (1st Dataset)	Visual	Vis+Attr
Discoid Eczema	80%	100%
Infantile Acne	60%	100%
Scabies	33.33%	93%
Total	57.78%	97.66%
Class (2nd Dataset)	**Visual**	**Vis+Attr**
Allergic Vasculitis	59.52%	100%
Atopic Eczema	48.67%	89.38%
Bullous Pemphigoid	56.09%	100%
Lichen Planus	38.09%	95.23%
Mycosis Fungoides	20.75%	96.22%
S. Cell Carcinoma	67.74%	98.38%
S. S. Melanoma	91.6%	98.6%
Total	61.09%	96.16%

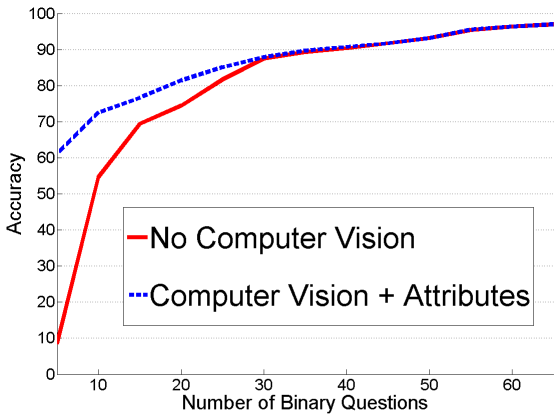


Figure 2: Results when no CV is involved. Although, CV and CV+Attributes results merge after a number of questions, CV increases accuracy without troubling the user with too many questions.

5 Conclusions

It is believed that there are between 1000 to 2000 skin conditions, and about 20% are difficult to diagnose. In the UK typical general practitioners receive minimal dermatology training. Our promising results from non-medical experts illustrate the potential clinical application of our work for health care providers, and also for places where access to health services are scarce.

We believe that we have for the first time applied a human in the loop visual recognition technique to diagnosis of skin diseases from visual images of affected areas. We have shown for some of the conditions, computer vision technique performs very poorly (as low as 20%), whilst human in the loop technique boosts the recognition rate to over 96%. Our future work is to apply the technique to a larger number of diseases, to refine the "Question and Answer" bank, and to implement the work on a smart mobile phone.

Acknowledgement: This work is partially supported by an EPSRC grant (EP/J501499/1).

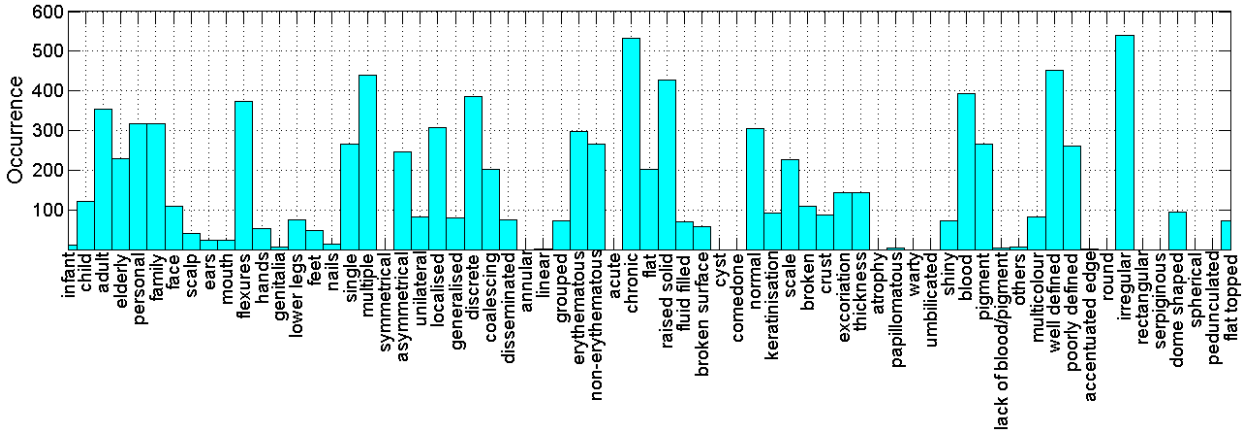


Figure 3: Frequency of answers picked by users in the 2nd dataset

References

[1] J Pfitzner M O'Rourke N Knight A Green, N Martin. Computer image analysis in the diagnosis of melanoma. *Journal of the American Academy of Dermatology*, 1994.

[2] R. Ashton and B. Leppard. *Differential diagnosis in dermatology*. Radcliffe, 2005.

[3] Lucia Ballerini, Xiang Li, Robert B Fisher, Ben Aldridge, and Jonathan Rees. Content-based image retrieval of skin lesions by evolutionary feature synthesis. *Applications of Evolutionary Computation*, 2010.

[4] Steve Branson, Catherine Wah, Florian Schroff, Boris Babenko, Peter Welinder, Pietro Perona, and Serge Belongie. Visual recognition with humans in the loop. In *Proceedings of the 11th European conference on Computer vision: Part IV*.

[5] Ela Claridge, Symon Cotton, Per Hall, and Marc Moncrieff. From colour to tissue histology: Physics based interpretation of images of pigmented skin lesions. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2002*.

[6] H. Ganster, P. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler. Automated melanoma recognition. *Medical Imaging, IEEE Transactions on*, 2001.

[7] Henning Muller, Antoine Rosset, Jean-Paul Vallee, and Antoine Geissbuhler. Integrating content-based visual access methods into a medical case database. In *Proceedings of the Medical Informatics Europe Conference (MIE 2003)*.

[8] Aude Oliva and Antonio Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int. J. Comput. Vision*, 2001.

[9] Francesco Orabona, Jie Luo, and Barbara Caputo. Online-batch strongly convex multi kernel learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.

[10] G. Qiu. Indexing chromatic and achromatic patterns for content-based colour image retrieval. *Pattern Recognition*, 2002.

[11] Guoping Qiu and Pong C. Yuen. Editorial: Interactive imaging and vision-ideas, algorithms and applications. *Pattern Recognition*, 2010.

[12] Pietro Rubegni, Gabriele Cevenini, Marco Burroni, Roberto Perotti, Giordana Dell'Eva, Paolo Sbano, Clelia Miracco, Pietro Luzi, Piero Tosi, Paolo Barbini, and Lucio Andreassi. Automated diagnosis of pigmented skin lesions. *International Journal of Cancer*, 2002.

[13] L Savolainen, J Kontinen, E Alatalo, J Rning, and A Oikarinen. Comparison of actual psoriasis surface area and the psoriasis area and severity index by the human eye and machine vision methods in following the treatment of psoriasis. *Acta dermatovenereologica*, 1998.

[14] A. Vedaldi and B. Fulkerson. Vlfeat: An open and portable library of computer vision algorithms. http://www.vlfeat.org/, 2008.

[15] A. Vedaldi, V. Gulshan, M. Varma, and A. Zisserman. Multiple kernels for object detection. In *Proceedings of the International Conference on Computer Vision (ICCV)*.

Performance Comparison of Low-level Vessel Detection Algorithms for Segmentation of X-ray Angiograms

Daniel Paredes Soto[1]
el10dap@sheffield.ac.uk
Philip Chan[2]
p.chan@sheffield.ac.uk
Peter Rockett[1]
p.rockett@sheffield.ac.uk

[1] Department of Electronic and Electrical Engineering
University of Sheffield
Sheffield
S1 3JD, UK

[2] Sheffield Vascular Institute
Northern General Hospital
Sheffield
S5 7AU, UK

Abstract

Using examples generated with a realistic model of the imaging physics, we have assessed and compared the performance of three popular multi-scale vessel detectors due to: Frangi *et al*., Sato *et al*. and Lorenz *et al*. We find that there is only a small difference between the performances of the Frangi and Sato detectors for the conditions considered in this work; for small-diameter vessels the response of these two detectors is complex although both suggest some promise for segmenting the vessels in the collateral bed around an occlusion. The Lorenz detector is shown to be the worst performing across a range of conditions.

1 Introduction

Occlusive vascular disease affecting arterial circulations is the major and fastest growing health problem worldwide, and underlies common conditions such as heart attack, stroke and peripheral vascular disease. The World Heath Organisation estimates that these diseases were responsible for 17.3 million deaths worldwide in 2008, forecast to rise to 23.6 million in 2030. Occlusion of major arteries is naturally compensated, to some extent, by the development of minor channels to carry blood around the occlusion, termed *collateral* circulation.

Traditionally, occlusive disease is treated by surgical methods although recently, pharmacological treatments, especially biological products have been researched. Unlike bypass or angioplasty, pharmacological candidates do not address the occlusion directly but are thought to encourage the formation of additional or better collateral vessels either through *angiogenesis* (the development of new blood vessels) and *arteriogenesis* (the development of flow-carrying arterial circulation from capillary precursors). Currently the success of treatments is judged on clinical endpoints like improved exercise ability or fewer amputations; such endpoints are massively confounded by other variables apart from blood flow changes.

Thus a major roadblock to the clinical evaluation of new drug treatments is the lack of an objective anatomical measurement of any enhancement in the blood conveying capacity of a collateral circulation—the development of such a methodology is our longer term goal. Such a direct measure of anatomical change from before to after-treatment would not be subject to confounding factors, and provide credible evidence of effect.

As a precursor to progress in this area, a means of accurately quantifying the properties the collateral network is needed and a fundamental step in this is the extraction of the, often fine, vessel segments which comprise the collateral bed. We have consequently studied the performance of a number of popular, low-level multi-scale vessel extraction algorithms [2, 4, 5] for digital-subtraction X-ray angiography, a common diagnostic procedure for patients presenting with the symptoms of occlusive vascular disease. This preliminary study seeks to compare performance in the ideal case of an isolated, straight vessel of infinite extent—see, however, Section 4 (Future Work). (We envisage that the low-level detector outputs will ultimately be combined to extract the whole vessel network using tracking, or Markov random field approaches; such higher-level methods are not the immediate concern of this paper.)

Here, we principally address two research questions: i) Which low-level vessel detector has the 'best' properties, and ii) How do these detectors perform on small-diameter vessels since these are a major concern for imaging collateral beds. We find that within the limits of this work, there is little to choose between the Frangi and Sato detectors. As far as we are aware, we are the first to present a *quantitative* comparison of these vessel detectors.

2 Methodology

We compare the performance of three much-used multi-scale detectors due to: Frangi *et al.* [2], Sato *et al.* [5] and Lorenz *et al.* [4]. All three algorithms calculate the eigen-decomposition of the Hessian matrix of the image at multiple scales. They then compute some scalar measure of 'vessel-ness' at each scale and select as the final vessel-ness measure, the maximum response across all scales. Essentially, the detectors only differ in how the hand-crafted scalar measures are calculated from the eigenvalues of the Hessian, λ_1 and λ_2. The Frangi detector [2] computes as a measure of eigenvalue eccentricity $\mathscr{R}_B = \lambda_1/\lambda_2$, where $|\lambda_1| \leq |\lambda_2|$. For some scale σ, the measure of 'vessel-ness', L_F is given by:

$$L_F(\sigma) = \begin{cases} 0 & \lambda_2 > 0 \\ \exp(-\mathscr{R}_B{}^2/2\beta^2)[1 - exp(-\mathscr{S}^2/2c^2)] & \text{otherwise} \end{cases} \quad (1)$$

where $\mathscr{S} = \sqrt{\lambda_1^2 + \lambda_2^2}$. We have used Frangi's suggestions of taking $\beta = 0.5$ and adjusting c for best results.

The vessel-ness measure due to Sato [5] is defined, for $|\lambda_1| < |\lambda_2|$, by:

$$L_S(\sigma) = \begin{cases} -\lambda_2 \exp(-\frac{\lambda_1^2}{2(\alpha_1\lambda_c)^2}) & (\lambda_2 < 0) \wedge (\lambda_1 < 0) \\ -\lambda_2 \exp(-\frac{\lambda_1^2}{2(\alpha_2\lambda_c)^2}) & (\lambda_2 < 0) \wedge (\lambda_1 > 0) \end{cases}$$

where $\alpha_1 < \alpha_2$. We have used Sato's suggested values of $\alpha_1 = 0.5$ and $\alpha_2 = 2$.

For the (parameter-less) detector of Lorenz [4] and $|\lambda_1| \leq |\lambda_2|$, $L_L(\sigma) = |\lambda_1|/|\lambda_2|$.

We compute the Hessian matrix over the ten linearly-spaced scales of $\sigma = 1$ to 10 pixel units and take the overall vessel-ness measure as the maximum detector response across all scales. We then threshold this measure to decide a label (vessel of non-vessel).

Obtaining ground-truth data in medical image processing is an enduring problem due to the uncertainties of inter- and intra-expert hand labelling. Here we adopt the well-established procedure of using synthetic data based on a physically-realistic model of the image formation process. Starting with out assumption of an isolated, straight vessel of infinite extent and circular cross-section irradiated with a uniform X-ray beam normal to the longitudinal vessel direction. We assume the vessel is uniformly filled with contrast agent. As the rays pass through the vessel, the intensity of the X-ray beam will be reduced according to the Beer-Lambert Law, $I = I_0 \exp(-\alpha\ell)$ where I_0 is the incident beam intensity, α is the absorption coefficient, and ℓ is the path length through the (contrast agent in the) vessel. See Figure 1. Cleary $\ell = \ell(x)$, where x is the spatial dimension in Figure 1, which also shows the resulting intensity profile projected onto the camera's focal plane. We selected α to give maximum contrast at the largest vessel diameter considered while allowing $\pm 3\sigma$ 'headroom' for the largest noise perturbations of the pixel values in the output image. The process thus mirrors a radiographer optimising the X-ray imaging conditions.

A distinguishing feature of the present work is the realistic image formation model: previous work, for example Krissian *et al.* [3], has usually assumed the projected vessel intensity profile is Gaussian with a physically-implausible infinite region of support. Drechsler and Laura [1] used the data of Krissian *et al.* in a qualitative comparison of vessel detectors.

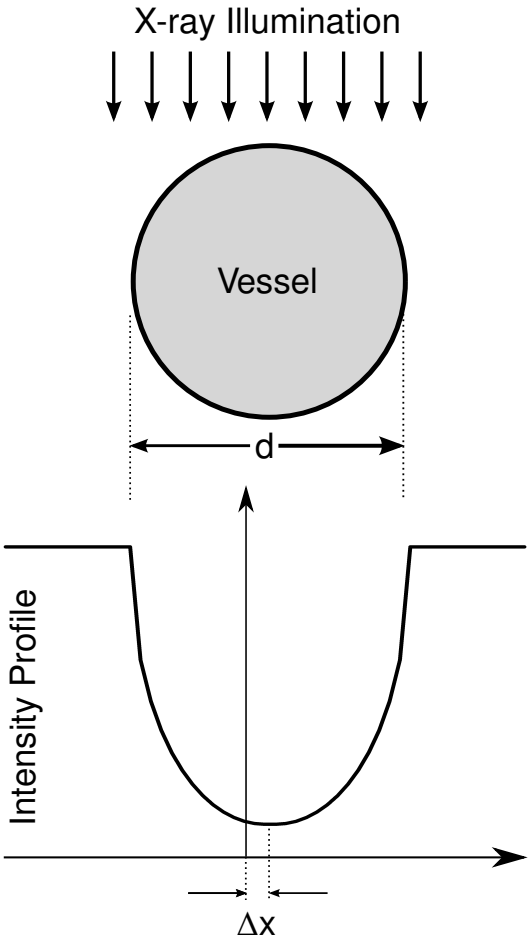


Figure 1: Formation of the vessel intensity model. The lower figure shows the projected intensity profile.

We have no *a priori* reason to suppose that a vessel will present at any particular orientation to, or displacement from the image lattice so we select a uniformly-random orientation $\theta \in [0 \ldots 45^o]$ and a displacement $\Delta x \in [-0.5 \text{ pixel} \ldots +0.5 \text{ pixel}]$ from the notional origin in Figure 1. (In practice, these are applied by an affine transformation of the projected intensity image above.) To account for the limitations of the imaging optics, we convolve the affine-transformed image with a Gaussian of $\sigma_{PSF} = 1$ and then add Gaussian-distributed noise of some σ_N which assumes that the dominant noise process is due to thermal noise in the camera/electronics. Finally, we quantise the pixel intensities into the range $[0 \ldots 255]$ to mimic analogue-to-digital conversion in the camera. This final image containing a single vessel is passed to one of the above detectors. In assessing a detector, we consider only the label of the central pixel

of the image through which the centreline of the vessel passes. Our rationale here is that the signal-to-noise ratio is maximised at the centre of the vessel—if the vessel cannot be detected at its centre, it cannot probably be detected anywhere else across its diameter. We thus implicitly investigate the upper bound of the detectors' performance. Additionally, each image represents an independent Monte Carlo trial.

To form counter-example images (*i.e.* genuine absence of a vessel) we form an image of uniform background intensity but with each pixel corrupted by noise of the same variance as the positive vessel examples. The method of generating positive and negative examples thus closely mirrors the process of digital-subtraction X-ray angiography.

We have repeated the above steps of generating positive and negative vessel examples in a Monte Carlo experiment (1000 trials) and and labelled with each of the detectors using various thresholds to construct receiver operating characteristic (ROC) plots. We also use the area-under-the-curve (AUC) as a summary statistic for the ROC plot.

3 Results

We have considered a range of vessel diameters ($D \in [1,3,5,7\ldots,15]$ pixel units) and noise powers ($\sigma_N \in [5,10,15,20,25]$). For brevity we only show results for the four permutations of small/large vessels, and lowest/highest noise powers since these represent the extremes of detector performance between which there is a more-or-less smooth variation. For the (physically unrealistic) absence of noise ($\sigma_N = 0$) all three detectors perform 'perfectly' (AUC = 1.0) apart from the case of unit vessel width ($D = 1$) where the Frangi and Lorenz detectors are tied (AUC = 0.95) and perform slightly worse than Sato (AUC = 0.97).

Fig. 2(a) and 2(b) show the ROC plots for $D = 1$ and the extremes of (a) low-noise ($\sigma_N = 5$) and (b) high-noise ($\sigma_N = 25$). Interestingly, in the low-noise regime all three detectors perform to some degree although the Sato and Frangi detectors are clearly better than Lorenz. (The straight line segments of the ROC plots are due to the discontinuities introduced by the case of $\lambda_2 = 0$ in Eqn. 1, and so on.) The detectors' behaviour is quite complex with the ROC plots crossing although there is generally little to choose between Frangi and Sato. Note that for the case of high noise, the performance of all three methods reduces to little better than random guessing—the 45^o line—as might be expected since the signal-to-noise (SNR) ratio becomes very low for this case. Since we have use a fixed value of attenuation coefficient (α), the SNR reduces as a function of reducing vessel diameter, hence small vessels are intrinsically harder to label.

For $D = 1$ and 3, and $\sigma_N = 25$ (high noise), there is little to choose between the performances of all three detectors. There is , however, an interesting transition between $D = 3$ and $D = 5$ (see Fig. 2(c) and (d)) where the Frangi and Sato algorithms begin to out-perform Lorenz, with Frangi slightly ahead of Sato, a trend which continues up to large diameters.

Fig. 2(e) and (f) show the corresponding plots for $D = 15$. At this stage, the Sato and Frangi algorithms both perform 'perfectly' (AUC = 1) regardless of noise power. The Lorenz detector, however, is clearly inferior, even for low noise powers.

As a summary, Fig. 3 shows the AUC statistic *vs.* vessel diameter D. The region between these two pairs of curves (low noise and high noise) delineates the useful operating envelope of the two detectors.

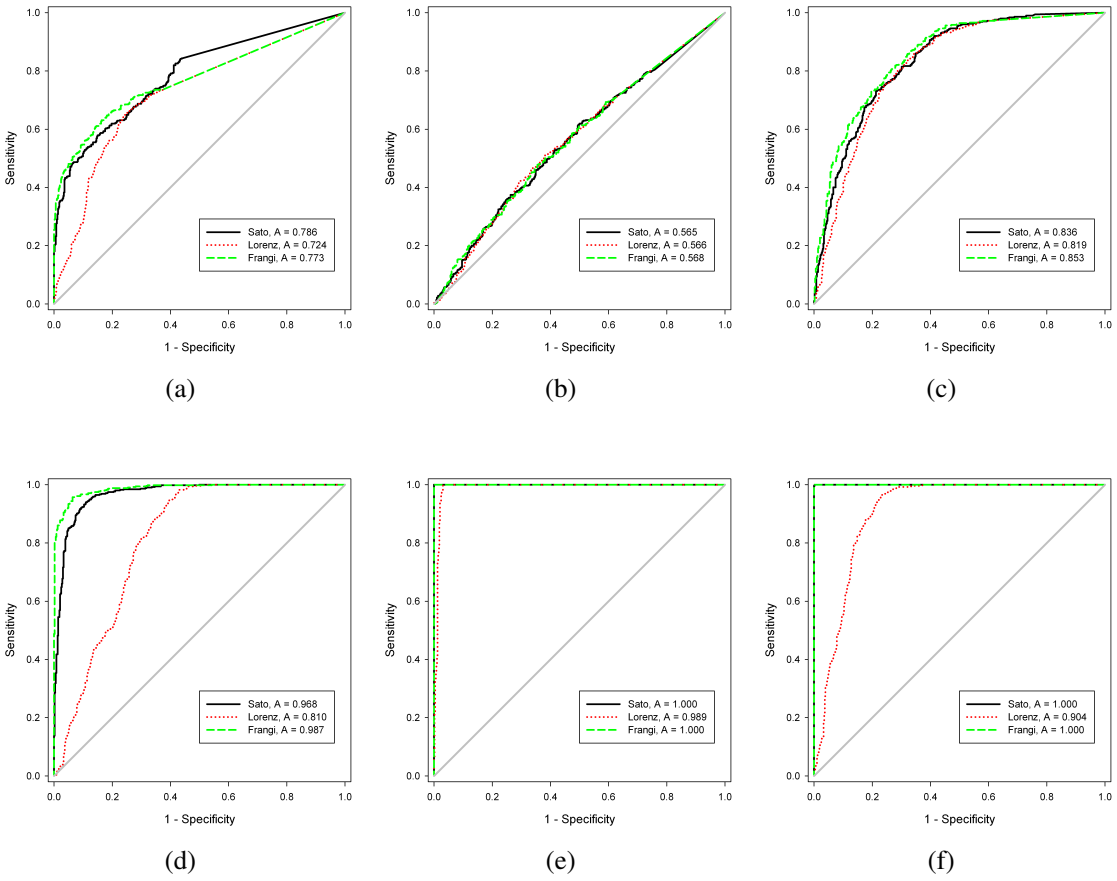


Figure 2: ROC plots for: $D = 1$ and (a) low-noise ($\sigma_N = 5$) and (b) high-noise ($\sigma_N = 25$). (c) $D = 3$ and (d) $D = 5$; high-noise ($\sigma_N = 25$). $D = 15$ and low-noise (e) ($\sigma_N = 5$) and (f) high-noise ($\sigma_N = 25$).

4 Future Work

The present preliminary report considers only the idealised case of a straight, isolated vessel. In future work, we will consider the effects on detector response of curved vessels and bifurcations. In addition, since the detectors are all multi-scale approaches, we anticipate that the responses from two vessels which approach closer than the scale of a Gaussian filter will interact with each other. For extracting vessel *networks* this is an important factor. Moreover, the vessels in a collateral bed frequently present as a 'corkscrew' and it is possible that the response (even) from an isolated corkscrew collateral could be affected by the proximity of other sections of the same vessel. These and other factors will be the subject of future studies.

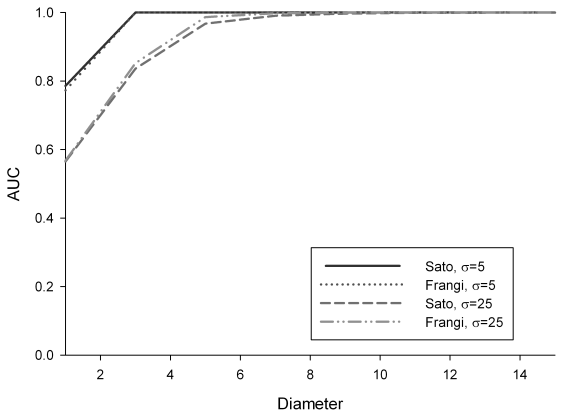


Figure 3: AUC *vs.* vessel diameter D for the Frangi and Sato detectors only.

We are assuming all vessels lie in a plane and can be imaged normal to the longitudinal

vessel direction. In reality, it is highly likely that vessels will have some component passing *into* a 2D image plane so extracting an entire network will be problematic. In this situation, 3D magnetic resonance angiography (MRA) will be highly desirable. We believe, given a suitable MRI simulator, the methodology in this paper can be directly applied to MRA images—this too is an area of future work.

5 Conclusions

From constructing a physics-based model of the process by which the image of a vessel is formed in digital-subtraction X-ray angiography, we have compared the performances of the vessel detectors devised by: Frangi *et al*., Sato *et al*. and Lorenz *et al*. On the basis of the present results, the Lorenz detector is clearly inferior to the Frangi and Sato detectors between which there is little to choose for anything other than the smallest vessel diameters. Whether Frangi or Sato is better for vessels on the 1 pixel scale would seem to depend on exact operating point. Nonetheless, these methods seem to offer some promise for segmenting fine vessels if combined with a contextual approach (*e.g*. Markov random fields).

The present results are limited to straight, isolated vessels. Whether any detector is superior for real, rather than idealised, vessel networks remains to be established in future work. For example, *qualitative* observations suggest the Frangi detector does not perform well at vessel bifurcations [1].

References

[1] K. Drechsler and C. O. Laura. Comparison of vesselness functions for multiscale analysis of the liver vasculature. In 10^{th} *IEEE International Conference on Information Technology and Applications in Biomedicine*, pages 1–5, Corfu, Greece, 2010.

[2] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In W M Wells, A Colchester, and S L Delp, editors, *Medical Image Computing and Computer-Assisted Intervention*, pages 130–137, Cambridge, MA, 1998.

[3] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trousset. Model-based detection of tubular structures in 3D images. *Computer Vision and Image Understanding*, 80 (2):130–171, November 2000.

[4] C. Lorenz, I.-C. Carlsen, T. M. Buzug, C. Fassnacht, and J. Weese. Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images. In 1^{st} *Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine and Medial Robotics and Computer-Assisted Surgery*, pages 233–242, Grenoble, France, 1997.

[5] Y. Sato, S. Nakajima, H. Atsumi, T. Koller, G. Gerig, S. Yoshida, and R. Kikinis. 3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. In 1^{st} *Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine and Medial Robotics and Computer-Assisted Surgery*, pages 213–222, Grenoble, France, 1997.

Ice-Water Phantom Localisation for Diffusion Calibration

Hossein Ragheb[1]
hossein.ragheb@manchester.ac.uk
Neil A. Thacker[2]
http://www.tina-vision.net/~nat/
David M. Morris[2]
David.M.Morris@manchester.ac.uk
Naomi H. M. Hogg[3]
Naomi.Hogg@icr.ac.uk
Alan Jackson[1]
alan.jackson@manchester.ac.uk

[1] The Wolfson Molecular Imaging Centre, Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK
[2] Imaging Sciences, Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK
[3] Institute of Cancer Research, Cotswold Road, Sutton, Surrey, UK

Abstract

When multiple scanners are used to acquire diffusion MR image data from the same patient, there is no guarantee that identical settings will result in comparable images of a specific organ and region. Despite this, for clinical use, any software developed to compute the apparent diffusion coefficient (ADC) must be expected to give equivalent results. What is needed is a phantom study with appropriate design which supports a calibration, so that appropriate settings for equivalent diffusion measurements can be defined among different scanners. Ideally the calibration process needs to be fully automatic, so that it can be used by non-experts in a clinical setting. We intend to develop software which automatically locates five cylinders in an ice water phantom and then measures the different diffusion values in the cylinders. The location algorithm uses an object recognition process which culminates in robust Likelihood estimation of position and orientation, computed using probabilistic Hough Transforms. This process estimates cylinder locations to an accuracy of a few pixels, even in the presence of field inhomogeneity and significant spatial distortion. Our assessment of performance includes quantification of possible errors due to; data inaccuracy, distortion due to poor shimming, signal to noise, field inhomogeneity and image clutter (i.e. ice). Results indicate that reliable localisation can be obtained using these methods for realistic clinical settings.

1 Introduction

Diffusion weighted (DW) magnetic resonance (MR) imaging has recently become a focus of attention for facilitating the diagnosis and treatment of tumors [1]. Specifically, to monitor the effect of new drugs administered to treat cancerous organs, DW-MR imaging is an attractive method compared to other alternatives including open surgery. However, there are many difficulties associated with the analysis of DW-MR images [2]. Recomendations have been made for tackling these problems, in order to ensure the validity of any corresponding

conclusions [2]. A common quantity that is used when analysing these images is the apparent diffusion coefficient (ADC). For consistency, it is important that any software developed to compute the ADC must give equivalent results on different scanners. One method to achieve this uses a non-biological phantom. For instance, Chenevert *et al*. [3] have recently performed a study using an ice-water calibration phantom and shown that it is possible to get good quantitative agreement between ADC's measured on different scanners. Gunter *et al*. [8] have used a phantom for the similar purpose of multi-centre MR (not DW-MR) scanner calibration but with a different design (and without using ice) and a different pattern recognition method.

This idea has recently been developed as part of the QUIC-CONCEPT project [1], providing us with two diffusion MR data sets. One data set corresponds to the standard shimmed scanner, while the other corresponds to a poorly shimmed scanner (Fig. 1). For diffusion measurement, image data are generally acquired at multiple b values, and exponential fits to this data generate diffusion parameters. In each data set there are five image slices for each of the four b values used (0, 100, 500 and 900), giving a total of 40 images with which to test our software. There are five cylinders of small diameter located inside the main large cylinder which is filled with ice and water. This maintains the temperature so that the measured diffusion is highly repeatable. One of the five cylinders is placed in the centre while the other four are symmetrically located around the centre (Fig. 1). Each of the five cylinders is filled with a different liquid which has specific diffusion characteristics. An automated calibration must first locate the five cylinders and the main topic of this paper is this task. We intend to use object recognition algorithms to identify and locate the phantom. However, the basic design of the phantom has 4-fold rotational symmetry, and this generates many combinatorial possibilities for location and orientation of the phantom which need to be combined and assessed. Probabilistic Hough transforms are used for this task, implemented as approximations to a likelihood estimation process, in order to deal with the large variations in the constraints on location parameters implied by the measured data.

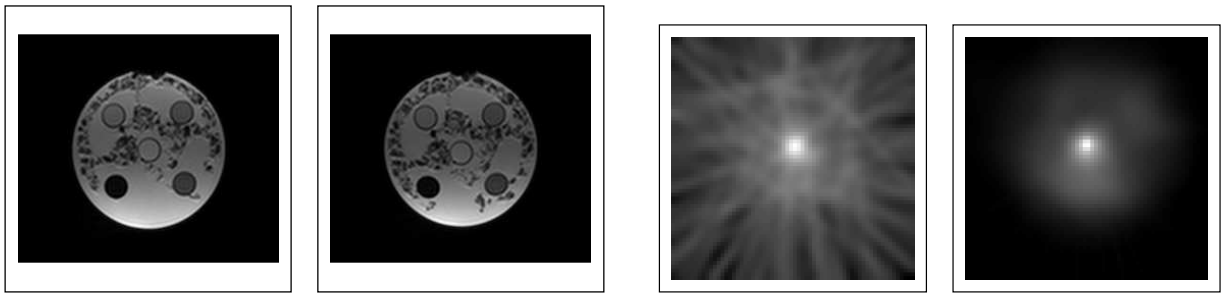


Figure 1: From left-to-right: example image slices from the standard shimmed (a) and poorly shimmed (b) phantom data; there are five different liquids with different diffusion characteristics inside the five cylinders, with ice and water at 0° temperature around them (ice appears as black irregularities); $L(x,y)$ for the standard (c) and $H(x,y)$ for the tuple Hough transfroms (d).

The combined process is expected to generate accurate estimates of phantom position even in the presence of image features such as inhomogeneity and ice, as well as being robust to significant spatial distortion and poor SNR. These issues are expected to compromise approaches based upon direct grey-level image matching and iterative optimisation.

[1]The QUIC-CONCEPT project has received funding from the European Union for the Innovative Medicine Initiative.

2 Methods and Results

We require a strategy to locate pixels falling inside each of the five cylinders in order to measure the diffusion of the liquid inside each. In our software, developed for this purpose, edge data is first extracted from the image and linked into extended curves. These curves are then approximated by polygonal approximations for convenience [4]. Each line in the polygonal description is then used as the focus for construction of a geometric histogram [5]. Geometric histograms are two dimensional and encode relative orientation and perpendicular distance between lines convolved with a PSF (point spread function) at a scale consistent with these measurements. The histograms can therefore be considered as density estimates of geometric co-occurrence. There are alternative versions of this basic representation which embody varying degrees of invariance. For the current work we used the most discriminative (least invariant) histogram, which considers two possible directions for each scene line separately, in order to compute a full 2π relative orientation. Two histograms are generated for each scene line over a range of possible scales.

The full set of histograms for all lines in the image are matched to a database of similar histograms generated using the model on the basis of a Bhattacharrya overlap [7]. The best match score and scale for each line is interpolated, so that each candidate match provides a constraint on the location, orientation and scale of the phantom in the image. The list of N best candidate matches for each scaled scene line is used as the basis for a series of probabilistic Hough transforms, which compute the robust log likelihood of centre, scale and orientation for the phantom. The 4-fold rotational symmetry of the phantom requires that $N >= 4$, and so we have used a value of $N = 8$ to guarantee that the histograms with high scores which are passed to the hough transform include the correct histogram. Likelihoods are computed on the basis of a perturbation model for the polygonalisation process, which assumes that errors on lines can be modelled using a constant uniform spherical error distribution σ_p on the end points [6]. These errors are propagated through to the parameter spaces in order to compute the required Hough entry distributions. The error model assumption is checked and the over-all scale of the spherical error determined using corresponding chi-squared distributions in data. A localisation error model with $\sigma_p = 2$ pixels was found to be approximately the correct value for this data.

Orientation and scale Hough transforms are computed as 1D histograms. The location Hough transform is 2D and entries are made for pairs of matched scene lines (n,m), i.e. a tuple Hough transform. For Gaussian density functions, the equivalent probabilistic form for this Hough transform $H(x,y)$ used to find the centroid of the phantom is given by

$$H(x,y) = \sum_m \sum_n log[p(x,y|n)p(x,y|m)] = \sum_m log[p(x,y|m)] \sum_n log[p(x,y|n)]$$

so that the total Hough entry can be considered as the square of the log Likelihood (i.e. standard Hough transform)

$$H(x,y) = L(x,y)^2 = (\sum_m log[p(x,y|m)])^2$$

In practice the 2D Gaussian distribution $p(x,y|n,m) = p(x,y|n)p(x,y|m)$ is computed directly using error propagation and then truncated at 3 S.D. to form a robust kernel. In comparison to the conventional Hough transform, the tuple based construction helps to remove background from the Hough array (Fig. 1). Cuts are placed on the allowable accuracy of each

pairwise constraint, as well as relative orientation at 5 S.D. of the propagated uncertainty in order to improve computational efficiency and further reduce background entries.

Estimated location, orientation and scale are the result of a global search of the parameter space (followed by a quadratic peak fit), thereby avoiding issues associated with local minima. These parameters are then available to project the model into the image in order to predict the location of each cylinder.

2.1 Evaluation of Cylinder Localisation

As a simple approach, we assess how well our algorithms can locate the cylinders for all image slices provided across all b values (40 images in total). This can be done by measuring the largest pixel shift between predicted and observed cylinder edges. Once this value is known it is possible to predict the quantity of usable pixels which might be reliably found in a circle around the predicted cylinder centre. Pixels falling inside these circles, may then be used to compute an average diffusion measurement. Fig. 2 shows the model overlayed on the image slices of Fig. 1. Here for the standard shimmed phantom a good match is achieved while for the poorly shimmed there is a slight mis-match. Here, it is clear that there is a systematic clockwise rotation of the model relative to the data. As the centroid of the phantom is easily estimated, the main localisation error for undistorted image geometry is generally of this nature, allowing us to attribute a sign (clockwise or anti-clockwise) to the residual error.

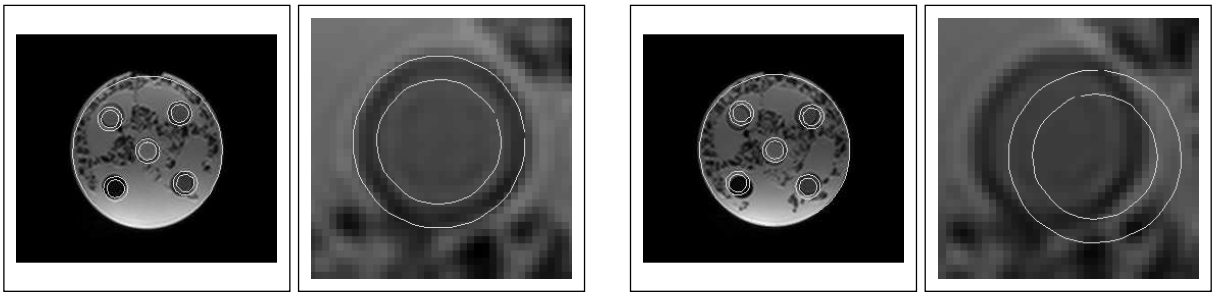


Figure 2: Locating all cylinders on image slices from standard shimmed (left) and poorly shimmed (right) data (Fig. 1) using our software; the display window is shown before and after zooming in to count the maximum out-of-fit pixels; the main error in localisation is generally in the form of a rotation; the maximum observable rotation error can be given a sign (positive if clockwise) in order to help identify systematic rotations across multiple images.

Tabulated results in Table 1 show that the mean shift of the model is comparable for both data sets, though there is evidence of a systematic rotation for the poorly shimmed data, due to significant systematic spatial distortion (columns 2-3). In both cases however, the mean maximum observed shift is ≈ 3.5 pixels for an internal cylinder diameter of 18 pixels (i.e. 20%). This error does not seem to come from problems of statistical stability (which is expected to be much better than one pixel), but rather spatial distortion of the original image. However, this does not mean that the achieved fit is distortion limited as there is also slight error on rotation due to noise and ice clutter.

In Table 1 (columns 4-5), we also list the percentage of field inhomogeneity based on the maximum coil correction together with the mean (first number) and the standard deviation (second number) for same image slices of different b values (every four rows). The degree of this inhomogeneity is substantial (a factor of two across the phantom) but does not appear to produce significant localisation errors. This is expected as the localisation is based

upon edges not absolute grey-level values. Further, in Table 1 (columns 6-7), we tabulate the percentage of ice in each image slice. These values are computed from the $b = 900$, via a manually controlled thresholding process following inhomogeneity correction, to an accuracy of around 5%. These typical percentages show that there is no observable correlation between the amount of ice and the inaccuracies seen in model matching (columns 2-3). We notice that these percentages are slightly larger for the case of poorly shimmed images.

image slices(b)	shift std	shift por	inhomogeneity (std)%	inhomogeneity (por)%	ice std%	ice por%	SNR std%	SNR por%
01(0)	+2	+3	100	140			1.45	1.76
06(100)	+2	+2	142	129			1.69	1.75
11(500)	+3	+4	125	137			1.41	1.66
16(900)	+2	+2	100	120	51.8	60.1	1.58	1.96
			116.7; 20.54	131.5; 8.96				
02(0)	-1	+2	100	143			1.66	1.93
07(100)	+4	+3	108	121			1.47	1.79
12(500)	+2	+2	100	100			1.59	1.75
17(900)	+3	+2	100	120	54.6	61.9	1.69	2.01
			102; 4.0	121; 17.57				
03(0)	+5	+2	76	106			1.70	2.20
08(100)	+2	+4	125	114			1.74	2.39
13(500)	-3	+4	100	100			1.67	2.00
18(900)	+1	+5	80	120	52.9	63.0	1.73	2.03
			95.25; 22.44	110; 8.79				
04(0)	+3	+2	114	120			1.81	1.98
09(100)	+3	+2	93	93			1.59	1.98
14(500)	-3	+4	78	80			1.74	2.27
19(900)	+3	+1	80	83	55.2	65.3	1.87	1.94
			91.25; 16.56	94; 18.2				
05(0)	-2	+3	100	100			1.81	2.20
10(100)	-2	+2	80	87			1.66	1.80
15(500)	+4	+2	78	80			1.83	1.95
20(900)	-3	+1	80	100	63.6	65.8	1.93	2.07
			84.5; 10.38	91.75; 9.94				

Table 1: Results are listed for both the standard shimmed (std) and poorly shimmed (por) image data sets; in columns 2-3 the positive sign for the number of pixels corresponds to the clockwise rotation while the negative sign corresponds to the anti-clockwise rotation; see text for further details.

Finally, in Table 1 (columns 8-9), we list the percentage of signal to noise ratios (SNR). Each ratio is computed from the distribution of second derivatives (for x and y) around zero, in a central rectangular region including all five cylinders, following coil correction. It is clear from the table that the SNR percentages are larger for the case of poorly shimmed data. However, it can be seen that for different b values the SNR changes by amounts which are small compared to measurement accuracy. Consequently, this level of noise is not large enough to test the robustness of the software to poor SNR. Hence, we selected an image with 50% ice and low geometrical distortion and added extra random noise until the software failed to locate the cylinders (where the midpoint of the predicted circles lies outside the

required liquid). It was found that for a factor of 2.5× more noise there is no observable degradation in performance, while the localisation algorithm fails 50% of the time at 3.4×, at which point the edge detector begins to fail to extract phantom structure. These noise levels are well beyond the expected levels for phantom data, consequently poor SNR is not expected to play a large part in the failure of localisation.

3 Conclusions

Our results suggest that the edge-based localisation method we have developed is robust to the main issues determining phantom data quality, such as field inhomogeneity, ice and SNR. In order to stay within the same range of image quality as seen in our tests we recommend that the quantity of ice is not allowed to rise above 70% and the SNR in the water to be kept above 5%. We have seen no evidence that the localisation process will not find the phantom in the 40 images provided with the standard parameters used for testing. Since there are five image slices per *b* value, it would be possible to check which of the five centres statistically agree (majority vote) in order to remove occasional localisation failures.

Future work now requires us to address two main issues. The main cause of localisation error is the spatial distortion which arises due to poor shimming. This cannot be accommodated by a localisation algorithm which assumes a rigid phantom model and implies that something must be done to accommodate slight shifts during the estimation of diffusion parameters. Also, the current algorithms localise the phantom with a 4-fold rotational ambiguity, due to the symmetry of the phantom. An additional marker has now been added to the phantom which we intend to use to select the correct rotation and as a quality check on predicted location.

References

[1] B. Turkbey, O. Aras, N. Karabulut, *et al*., Diffusion-Weighted MRI for Detecting and Monitoring Cancer: A Review of Current Applications in Body Imaging, *Diagnostic Interventional Radiology, Turkish Society of Raiology*, 18(1):46-59, 2012.

[2] A.R. Padhani, G. Liu, D. Mu-Koh, *et al*., Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations, *Neoplasia*, 11(2):102-125, 2009.

[3] T.L. Chenevert, C.J. Galban, M.K. Ivancevic, *et al*., Diffusion Coefficient Measurement Using a Temperature-Controlled Fluid for Quality Control in Multicenter Studies, *J. Magnetic Resonance Imaging*, 34:983-987, 2011.

[4] D.G. Lowe, Three-Dimensional Object Recognition from Two-dimensional Images, *Artificial Intelligence*, 31(3):355-395, 1987.

[5] A.C. Evans, N.A. Thacker and J.E.W. Mayhew, The Use of Geometric Histograms for Model-Based Object Recognition, *Proc. British Machine Vision Conference*, 429-438, 1993.

[6] A.P. Ashbrook, N.A. Thacker, P.I. Rockett and C.I. Brown, Robust Recognition of Scaled Shapes using Pairwise Geometric Histograms, *Proc. British Machine Vision Conference*, 503-512, 1995.

[7] N. A. Thacker, F. Ahearne and P. I. Rockett, The Bhattacharyya Metric as an Absolute Similarity Measure for Frequency Coded Data, *Kybernetika*, 34(4):363-368, 1997.

[8] J. L. Gunter, M. A. Bernstein, B. J. Borowski, *et al*. Measurement of MRI Scanner Performance with the ADNI Phantom, *MEDICAL PHYSICS*, 36(7):2193-2205, 2009.

Compensating for Drusen in Retinal Vessel Segmentation

Lei Zhang[1]
Lei.zhang@uea.ac.uk
Mark Fisher[1]
http://www2.cmp.uea.ac.uk/~mhf/
Wenjia Wang[1]
Wenjia.Wang@uea.ac.uk

[1] School of Computing Sciences, University of East Anglia, UK

Abstract

Segmenting vessels from retinal images, like segmentation in many other medical image domains, is a challenging task, as there is no unified way that can be adopted to extract the vessels accurately. This paper focuses on removing Drusen (isolated areas of brightness), a problem and a common source of error in automated vessel detection. We use the log-Gabor filter to calculate the local energy in the image and so identify the Drusen. A pre-processing stage comprising an average filter is then used to smooth the affected area in order to eliminate the imperfection before vessels are classified by an existing technique. The results confirm that the modified method shows some improvement compared to a previously published method when tested using the STARE database.

1 Introduction

In optometry, the appearance of blood vessels in the retina is an important indicator, which is routinely examined to detect diseases, including diabetes, hypertension, and glaucoma. For example, diabetic retinopathy (DR) is basically a main complication caused by diabetes, and it would lead to blindness if the diabetic patient did not receive the treatment in time. But the assessment of the retina is a skilled time consuming task [6], and as such it has been the focus of research into automatic assessment techniques. Amongst the methods or algorithms that have been presented, Echevarria and Miller [11] propose a method that utilizes the concept of level sets to remove the noise and a fast marching algorithm to trace the vessels. They also use a filter designed by Chaudhiri et al. [12] to enhance the image contrast. Farnell [4] presents a method that adopts a multi scale line operator (MSLO) filter to enhance the retinal blood vessels in the image and then implements Otsu's threshold selection approach [8] to segment the MSLO filter response. In [5] Diego Marín et al. adopts a neural network (NN) scheme and a 7-D vector comprising gray-level and moment invariant features is used for the pixel classification task. The general framework of blood vessel segmentation presented in [3] is similar to the above methods. Innovatively, Wu et al. [3] combines a diffusion enhancement Hessian-based, and matched filters, which incorporate edge constraints to produce a single

compound filter to enhance the image. This filter needs only one matched filter at each scale and is more efficient than other matched filters which require multiple filter applications per pixel [3]. Soares and Leandro et al. [7] proposed a scheme using the 2D Gabor wavelet to implement the noise filtering and vessels enhancement, then using a Bayesian classifier with class-conditional probability density functions named Gaussian mixture model (GMM) to identify whether pixels are vessel or not. Cinsdikici and Aydin [9] proposed a novel approach which is a hybrid model of matched filter (MF) and ant colony algorithm. The novel method overcomes an imperfection of classic MF that it improves the performance of detecting small branches of vessels (capillaries). Bob and Lin et al. [2] proposed a novel extension of the MF approach which is named MF-FDOG to distinguish the vessels from non-vessel step edges. The basis of MF-FDOG is that it detects the vessels objects by thresholding the retinal image's response to a zero- mean Gaussian function while using the local mean of the response to the first-order derivative of Gaussian (FDOG) to adjust the threshold level in order to remove the non-vessel edges. Although, many methods have been proposed and some success has been achieved, the algorithms frequently fail due to the challenging nature of retinal image analysis. The images are often contaminated by noise and suffer low contrast between the vessels and surrounding background. Cases in which suffer from existing damage due to disease can be particularly problematic and our work is focused in this area. In these pathologies local small brightness blobs especially located round vessels have the most important influences for segmentation. Hence, we aim to find what the brightness spots are and how we can reduce their affects for vessels segmentation.

1.1 Brightness area (Drusen)

Drusen (e.g. illustrated in the figure1) are deposits of epithelial cell waste that are located beneath the retina. They present as yellowish blobs which are brighter than the orange background, hence it appears more brightly in the green band of retinal images. This phenomenon may affect the accuracy of vessel segmentation which we found in the experiments (Fig. 2 b) and indicated in the other literatures [2][8]. Our scenario is that removing the brightness blobs (like Drusen) before implementing the segmentation in order to reduce the Drusen caused influences of vessel segmentation.

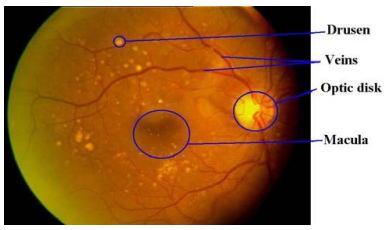


Figure 1 colour retinal image sample [13]

1.2 Materials

The experimental data，we choose from the STARE dataset which is publicly available on the website www.ces.clemson.edu/ahoover/stare/. It's originally collected by Hoover et al. [1]. The images of dataset were stored as PPM format, 8 bits per colour channel and digitalized to size 700*605. The dataset contains two manual segmentations results made by 2 observers， and the first set of manual segmentation results is used as ground truth to evaluate the performances of two methods in our experiment.

2 Method

The framework of our method is primarily composed of four stages: 1. pre-processing, 2. Drusen detecting and removing, 3. using Matched filter to generate the maximal filter responses and 4. Convert the responses to binary by local entropy thresholding and using length filtering to remove the isolate objects.

2.1 Pre-processing

In pre-processing stage, we convert colour retinal images into gray level images by isolating the green channel, since other authors (e.g. Wu et al. [3]) have noted that contrast between vessels and background is enhanced in this channel. A two dimensional Gaussian filter is then adopted to remove the vessel reflection and smooth the image. From experiment experience we choose the parameters as σ_x = 1.5 σ_y=2. At the end of pre-processing stage, we adjust the image to enhance the contrast between the vessels and background.

2.2 Drusen detection using local energy

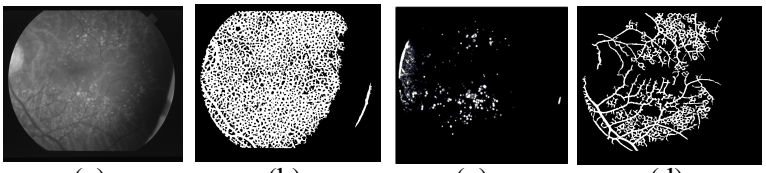
A texture-based Drusen detection method [13] is adopted in our experiment, the texture of the drusen can be characterised in terms of local energy. The Local energy has been defined in [10] as the sum of squared responses of orthogonal pairs of Gabor or log-Gabor filters. Compared to Garbor filters, the log-Gabor filter has several advantages. It allows larger bandwidths to be efficiently implemented (typically 1 to 3 octaves), which in turn allows one to capture more features [13]. A log-Gabor filter can be expressed as

$$G(f,\theta) = \exp\left\{\frac{-[\log(f/f_o)]}{2[\log(\sigma_f/f_o)]}\right\} \times \exp\left\{\frac{-[\theta-\theta_o]}{2\sigma_\theta^2}\right\} \qquad (1)$$

Where the former part of formula is radial component and the latter part is orientational component. f_o is the central radial frequency, θ_o is the filter direction. σ_θ, which defines the angular bandwidths, and σ_f represent the radial bandwidth which is computed by both σ_f and f_o. The local energy at every pixel(x,y) is calculated by summing squares of even and odd symmetric log-Gabor filter responses at every point (x,y). It is obtained as

$$E_{\theta_O}^{f_o}(x,y) = (R_{\theta_O}^{f_o,even}(x,y))^2 + (R_{\theta_O}^{f_o,odd}(x,y))^2 \qquad (2)$$

Where $R_{\theta_O}^{f_o,even}(x,y)$ and $R_{\theta_O}^{f_o,odd}(x,y)$present the responses of even and odd symmetric log-Gabor filters respectively. An example of the energy map is illustrated in Fig. 2 (c). The brighter areas of the image indicate higher values of local energy.



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

Figure 2 drusen detection and results, (a) is gray level image (b) segmented result using method [14] (c) binary drusen mask (d) segmented result using our modified method

Once Drusen detected they are removed by an averaging filter applied to the areas identified by the local energy map. We choose a filter size of 40*40, and before apply the filter on the objective areas of image (the Drusen areas), we transform the local energy map to a binary image which is used to determine the range of averaging area. In our experiment, these areas, actually, are Drusen. The threshold is simply chosen by calculating the average value of energy map.

2.3 Matched filter

In [14] the authors adopt a 2-D Gaussian filter kernel with 12 different orientations. The vessel can be regarded as typical line feature with different orientations in the retinal image, and it may be oriented at any angle, hence, as noted in [14], we should implement filter kernels that can be rotated and aligned in any direction. In our experiment, we applied a set of 12 orientations (0°, 15°, 30°, 45°, 60°, 75°, 90°, 105°, 120°, 135°, 150°, 165°) of Gaussian kernel with one scale (σ =1.75). The size of each Gaussian kernel is 16*15 pixels. Fig. 3 illustrated the Gaussian kernel with 12 different orientations. The enhanced image then is generated by taking the maximum filter responses at each pixel.

Figure 3 Gaussian kernels with 12 different orientations.

2.4 Segmentation method

Considering dependences of intensities between each image pixel, we follow the same approach to Chanwimaluang et al. [14] and take into account the spatial distribution of gray level. The segmentation method is composed of two steps. Firstly, we use a local entropy thresholding algorithm to compute the optimal threshold for classifying vessels and background. The second step is using length filtering to remove the isolate objects that may be misclassified. The primary task in local entropy thresholding is that of choosing parameters for the calculation of entropy. An optimal threshold is chosen based on the maximum of sum of second derivative entropies on different gray levels.

3 Experiment results

The parameters of method are trained on images which don't have manual segmentation result (ground truth), the test and evaluation procedure are applied on 18 STARE images for both original method proposed by Chanwimaluang et al. [14] and our improved method. We use standard measures of performance, namely, sensitivity (TP/TP+FN), specificity (TN/TN+FN) and accuracy (TP+TN/TP+FP+TN+FN). The sensitivity and specificity are the factors which indicate success rate of classifying vessel and non-vessel structures, respectively. The sensitivity is also known as TPF which represents the ratio of the number of correctly classified vessel pixels to the number of total vessel pixels in the ground truth. The accuracy indicates the ratio of total well classified pixels. In addition, the

algorithm's performance was also measured with receiver operating characteristic (ROC) curves. The ROC curves for each image are obtained by manually thresholding the image with the threshold values starting from 0 to 1 in a step of 0.01 (e.g. Fig. 4). The areas under the ROC curves for both original method and our method are listed in Table 1. Az indicates the area under the ROC curves, also is known as AUC. Columns indicated by 1 and 2 are our modified method and Chanwimaluang et al. [14] method, respectively. The last row of Table 1 shows the average sensitivity, specificity and accuracy values.

Image	Sensitivity		Specificity		Accuracy		Az	
	1	2	1	2	1	2	1	2
Im0001	**0.700938**	**0.66857**	**0.9509**	**0.9279**	0.9310	0.9072	0.9081	0.8926
Im0002	**0.738628**	**0.725829**	**0.9200**	**0.9001**	0.9080	0.8885	0.9152	0.9065
Im0003	**0.769185**	**0.763385**	**0.9016**	**0.8957**	0.8937	0.8878	0.9136	0.9112
Im0004	**0.534416**	**0.574709**	**0.9840**	**0.9809**	0.9507	0.9508	0.9273	0.9181
Im0005	**0.582482**	**0.614554**	**0.9729**	**0.9672**	0.9377	0.9354	0.9323	0.9347
Im0044	**0.690704**	**0.861405**	**0.9736**	**0.9112**	**0.9539**	**0.8204**	0.9166	0.9406
Im0077	0.766949	0.804131	0.9637	0.9526	0.9479	0.9407	0.9314	0.9261
Im0081	0.824229	0.86459	0.9517	0.9382	0.9422	0.9327	0.9410	0.9346
Im0082	0.726719	0.792195	0.9766	0.9634	0.9569	0.9500	0.9331	0.9330
Im0139	0.708283	0.765393	0.9744	0.9465	0.9530	0.9320	0.9267	0.9275
Im0162	0.717064	0.75338	0.9721	0.9600	0.9539	0.9453	0.9492	0.9488
Im0163	0.745046	0.81283	0.9819	0.9669	0.9636	0.9549	0.9544	0.9519
Im0235	0.795867	0.81401	0.9495	0.9440	0.9358	0.9325	0.9239	0.9235
Im0236	0.697211	0.839155	0.9736	0.9310	0.9485	0.9227	0.9370	0.9366
Im0240	0.663206	0.739102	0.9679	0.9577	0.9367	0.9354	0.9264	0.9187
Im0255	0.720708	0.807852	0.9743	0.9584	0.9516	0.9449	0.9478	0.9474
Im0291	0.74232	0.772621	0.9817	0.9769	0.9696	0.9666	0.9576	0.9518
Im0319	0.747767	0.564538	0.9644	0.9851	0.9550	0.9670	0.9160	0.9107
Average	**0.7151**	**0.7521**	**0.9630**	**0.9480**	**0.9439**	**0.9286**	**0.9310**	**0.9286**

Table 1 Performance results on STARE data by both our method and Chanwimaluang et al. [14] .

Considering the sensitivity, Specificity and Accuracy with ROC curve, we can say that the modified method improves the segmentation performance compared to original method, particularly reduce the miss-segmentation rate (FPF) that the rate of tissue not belong to vessel are miss-segmented as vessels. Observing from the Fig. 4, in contrast to original method's plus (+symbols) points line, we can find that the red point curve is closer to the top left corner.

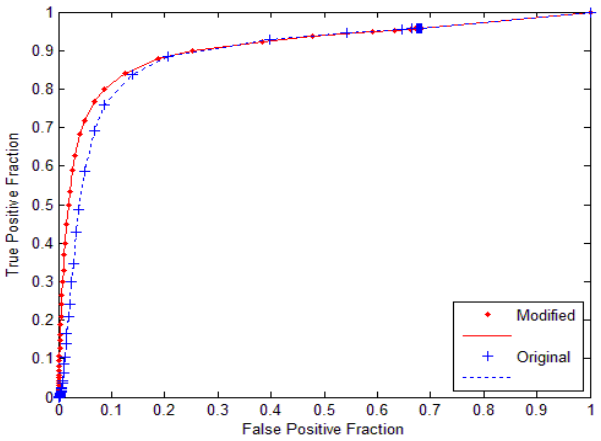


Figure 4 ROC curves for the first image of STARE

This confirms our method achieves some success in reducing the effects caused by Drusen which present as light spots in the image and produce false positives.

4 Conclusion

In this paper, we explored a method for extracting the vessel structure from retinal images. We use a log-Gabor filter to detect Drusen that produce bright areas in the image that can

produce errors in the vessel segmentation quality. Considering our experimental results (the sensitivity 0.7151, specificity 0.9630 and accuracy 0.9439 with AUC area under the ROC curve 0.9310) compare with original method (sensitivity 0.7521, specificity 0.9480 and accuracy 0.9286 with AUC area under the ROC curve 0.9286) we can say that the modified method improves the segmentation performance compare to original method, particularly reduce the miss-segmentation rate (FPF) that the rate of tissue not belong to vessel are miss-segmented as vessels. We can conclude that the additional step makes a valuable contribution and a significant improvement.

References

[1] A. Hoover, *et al.*, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response", *IEEE Trans. Med. Imaging* 19 (3) 203–210, 2000

[2] Bob Zhang, *et al.,* "Retinal vessel extraction by matched filter with first-order derivative of Gaussian", Computers in Biology and Medicine, 40(4), 438-445, 2010

[3] Changhua Wu, *et al.*, "A general framework for vessel segmentation in retinal images", In Proceedings of CIRA. 37-42, 2007.

[4] D.J.J. Farnell, "Initial Results of an Automatic Blood-Vessel Segmentation Procedure in Digital Fundus Photographs via Multiscale Line Operators and Global Threshold Selection." 2009. http://dircweb.king.ac.uk/miua2009/pdfs/miua/017.pdf

[5] Diego Marín, *et. Al.*, "A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features", *Medical Imaging, IEEE Transactions on ,* Volume:30, 146 – 158, 2011.

[6] H. Schneiderman, The Funduscopic Examination, http://www.ncbi.nlm.nih.gov/books/NBK221/, 1990.

[7] J. V. B. Soares, *et al.*, "Retinal vessel segmentation using the 2D Gabor wavelet and supervised classification," IEEE Trans. Med. Imag., vol. 25, no. 9, 1214–1222, 2006.

[8] L. Gang, O. Chutatape, and S. Krishnan, "Detection and measurement of retinal vessels in fundus images using amplitude modified second-order Gaussian filter," *IEEE Trans.Biomed. Eng.*, vol. 49, no. 2, 168–172, 2002.

[9] M. Cinsdikici and D. Aydin, "Detection of blood vessels in ophthalmoscope images using MF/ant (matched filter/ant colony) algorithm", Comput. Methods and Programs Biomed. 96, 85–95. 2009.

[10] M.C. Morrone and R. Owens, "Feature detection from local energy", Pattern Recognition Letters 1 103–113. 1987.

[11] P. Echevarria T. Miller J. O'Meara, Blood Vessel Segmentation in Retinal Images February 8, 2004. http://robots.stanford.edu/cs223b04/project_reports/P14.pdf

[12] S. Chaudhuri, *et al.*, "Detection of blood vessels in retinal images using two-dimensional matched filters". *IEEE Trans. on Medical Imaging*, 8(3):263-269, 1989.

[13] S. Garg, J. Sivaswamy and G. D. Joshi, "Automatic Drusen Detection from colour retina", http://cvit.iiit.ac.in/papers/saurabh06Automatic.pdf

[14] T. Chanwimaluang and Guoliang Fan. "An efficient algorithm for extraction of anatomical structures in retinal images", In Proc. of ICIP (1), 1093-1096, 2003

[15] T. Kurita, N. Otsu and N. Abdelmalik. "Maximum likelihood thresholding based on population mixture models," Pattern Recognition 25, pp. 1231–1240, 1992.

Coping with Noise in Ultrasound Images: A review

Jonathan Francis Roscoe
jjr6@aber.ac.uk
Hannah Dee
hmd1@aber.ac.uk
Reyer Zwiggelaar
rrz@aber.ac.uk

Department of Computer Science
Aberystwyth University
Wales, UK

Abstract

Ultrasound is notorious for having significant noise with a low signal-to-noise ratio. This inhibits the performance of segmentation and causes difficulty for clinical evaluation, thus noise reduction is paramount to achieving adequate segmentation in ultrasound images. Consequently, the modeling and handling of noise is a significant area of research. In this review paper we introduce the typical characteristics of noise in B-mode ultrasound and analyse th performance of multiple state-of-the-art methodologies for dealing with such noise. A similar paper was written by by Coupé et al. [8] when they introduced OBNLM; though we provide an independent review and generalised description of the problem area. We also discuss the issue of typical image quality assessment methods and consider the impact speckle noise could have on ultrasound image analysis.

Three state-of-the-art denoising algorithms (SRAD, SBF, and OBNLM) are evaluated using three different image quality assessment methods (SSIM, MSE and USDSAI) in comparison with traditional filters such as Lee's. We worked with simulated phantom images, as well as prostate ultrasound images to assess these methods. SRAD and OBNLM seem to be the most effective algorithms and in our discussion we contemplate ways in which they might be further expanded.

1 Introduction

Ultrasound is notorious for having pervasive noise with a low signal-to-noise ratio which inhibits the performance of segmentation algorithms and causes difficulty for clinical evaluation. Consequently, the modeling and handling of noise is a significant and continuing area of research. We are primarily concerned with speckle, due to its prevalence in ultrasound and the fact that it may be considered noise or a source of information.

2 Types of Noise and Their Cause

There are two basic models for noise behaviour [5]. The first, **additive** noise is generally more common. It is independent of image data - thermal noise and noise caused by quantisation are common examples. Secondly, **multiplicative** noise is related to image data and

is often found in coherent imaging systems (such as ultrasound) whilst uncommon in other modalities; the typical example is speckle noise which occurs due to variation in the surface being imaged. For the purpose of this paper, and ultrasound in general, we focus on speckle noise, though the most relevant varieties of noise are:

Gaussian noise can frequently be observed as the result of thermal agitation (a.k.a Johnson-Nyquist noise), film grain (sometimes modelled as Poisson noise) and photon counting. It is indicative of the physical characteristics of the imaging methodology. **Speckle** is present in coherent imaging systems; the backscatter waves from a surface may constructively or destructively interfere causing modulation in phase and amplitude observed as variation of high and low intensities known as speckle. It is therefore characteristic of the surface being imaged and inherently multiplicative. Speckle is a significant component of ultrasound and other noise may be considered negligible. **Quantisation noise** arises from the process of transforming continuous data into discrete values (quantisation) and is a mandatory component of digital acquisition. The result is a uniform degradation in resolution characterised by a blocky appearance. Sufficiently high resolution data acquisition can mitigate this.

Year	Author	Technique	Dataset Used
2009	Coupé et al. [8]	Optimised Bayesian NL-means (OBNLM)	2D intraoperative brain images and 3D liver images
2006	Yue et al. [33]	Non-linear Multiscale Wavelet Diffusion (NMWD)	Echocardiographic images
2006	Acton et al. [27]	Squeeze Box Filter (SBF)	Field II simulation
2002	Yu and Acton [32]	Speckle Reducing Anisotropic Diffusion (SRAD)	Carotid artery ultrasound images
2001	Achim et al. [2]	Bayesian Multiscale Estimator (wavelet)	Simulated speckled images
1990	Perona and Malik [22]	Anisotropic Diffusion	Not originally applied to ultrasound, but has been adapted by [32] and others
1989	Loupas et al. [18]	Adaptive weighted median filter (AWMF)	Various images, including liver and gallbladder
1987	Kuan et al. [17]	Non-linear MSE minimisation	Not originally applied to ultrasound
1980	Lee et al. [15]	MSE minimisation	Not originally applied to ultrasound

Table 1: Trends in ultrasound specific denoising

3 Noise Reduction

Ultrasound preprocessing typically involves a significant noise reduction stage to mitigate the presence of speckle and improve segmentation. Conventional image processing filters such as median [18], Lee's [15] and Kuan [17] have traditionally being used to this end. Current analysis of the efficacy of noise filters has been conducted by the original authors, and as such, we felt an independent assessment of major techniques would be of value to the community.

Table 1 provides an overview of both traditional and state-of-the-art filters. Recent developments have been based around anisotropic diffusion and non-local means. Anisotropic

diffusion techniques essentially convolve the image with a Gaussian kernel and are superior to traditional techniques at edge preservation. Non-local means (NL-means) [7] filters generate a weighted average based on similarity between pixel neighbourhoods; this prevents unncessary averaging with non-similar regions. NL-means is increasingly popular, and has been adapted to the CUDA platform for real-time ultrasound [9].

In the following subsections we present an overview of typical image quality metrics, as well as the results of performing vairous denoising algorithms on real and simulated ultrasound data.

3.1 Assessing Image Enhancement

A variety of metrics can be used in order to quantify the quality of signals and the efficacy of filters. A long favoured and reliable measure is the mean squared error (MSE) [3] and the related peak signal-to-noise ratio (PSNR) [13]. The output from these methods do not necessarily correlate with human perception of quality and results may be inconsistent across content, most significantly they do not discern structure in images [29]. Whilst all of these metrics are easily implemented and performed, they have particularly obtuse view on what makes a "good" quality image. MSE and PSNR have prevailed as a popular indicators of signal quality due to simplicity in implementation and ease of understanding.

Interest has grown in more sophisticated indicators of signal fidelity. The structural similarity index (SSIM) [30] is a metric targeted at quality assessment based on a reference image with consistency in structural information being an indicator of quality. A SSIM approaching 1 indicates similarity, whilst a SSIM approaching -1 represents dissimilarity, a value of 1 itself would occur in the event of identical images. One study has indicated a link between PSNR and SSIM [12], from which we infer MSE, PSNR or SSIM may be analogous and adequate for the majority of cases although some differences in sensitivity should be noted.

In the context of ultrasound imaging Tay et al. introduced a modified Fisher discriminant contrast metric referred to as the ultrasound despeckling assessment index (USDAI), a ‘large USDSAI would indicate that [the algorithm] produces desirable restoration or enhancement results’ [27], this metric is used to demonstrate the superiority of SBF to SRAD and traditional filters. Coupé et al. have used the same framework to demonstrate superiority of OBNLM to SRAD, SBF and conventional NL-means[8].

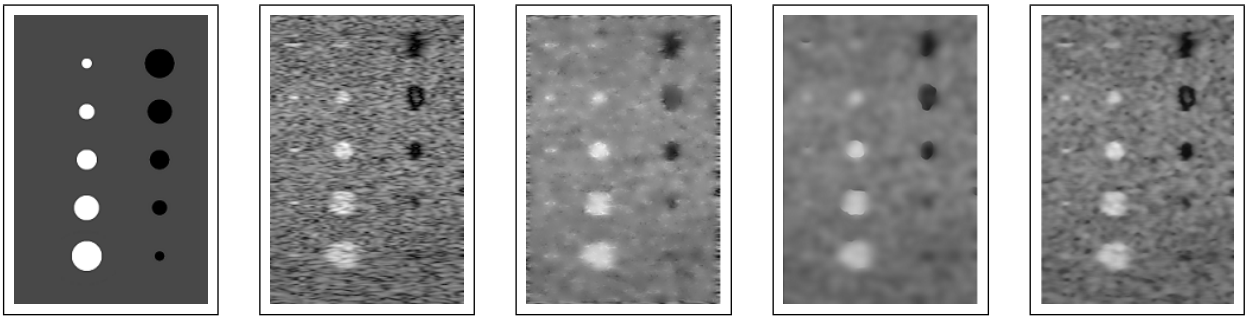
Although USDSAI is freely available, no independent evaluation has yet been carried out. Therefore we chose to compare and contrast the MSE, SSIM and USDSAI metrics. Perhaps in the future a well defined solution to the problem of image quality assessment will be available. USDSAI is a good step in this direction given its specificify to homogeneous classes whilst being sensitive to changes across regions.

4 Experiment and Results

Regardless of the evaluation framework, it is important to have a reference image to which image enhancements can be compared. As ultrasound data is of poor quality due to acquisition technology, the best source for such a data set is simulated phantom images. These images should contain structural features of different contrasts, for example Figure 1.

Many techniques for simulation of B-mode images have been published [14, 19, 23, 33]. Whilst these methods may vary in their ability to precisely mimic speckle, they do possess statistically significant characteristics that make them suitable. In Figure 1 we demonstrate

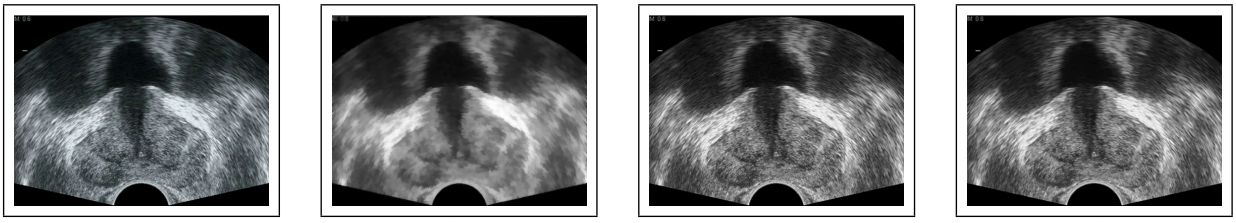
the visual appearance of various state-of-the-art denoising approaches applied to a Field [14] simulated ultrasound image (shown in Figure 1a). We compared four filters: SRAD. SBF, OBNLM and median. Each was run multiple times - with a variety of parameters. The effect of filters can be difficult to appreciate visually, and assessment was carried out using MSE, SSIM and USDSAI. The best results from each evaluation method (not limited to a single set of parameters) are listed in Table 2.



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

Figure 1: Comparison of denoising methods applied to phantom generated with Field II. These images all have the best USDAI score for the specific denoising method. (a) Template image used as ground truth (b) Field II simulated image. Effects of denoising with (c) SBF (d) SRAD(e) OBNLM.

Figure 2 shows the algorithms performed on a real prostate ultrasound image. It is extremely difficult to perform image enhancement quality assessment on clinical data as there is no true ground truth. The output of each algorithm at various parameters was compared with the original image using SSIM and MSE metrics to assess the amount of degradation resulting from each. It was not possible to use USDSAI for this adata due to the unkown classes present in the image. Interestingly, regardless of the distance between metric results; generally the best result using one metric will be the same for the other.



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

Figure 2: Comparison of denoising methods applied to real prostate ultrasound images. Note that the better algorithms (SRAD, OBNLM) preserve detail much more effectively. These images all have the best SSIM score for the specific denoising method. (a) Original (b) SBF MSE: 280, SSIM: 0.65(c) SRAD MSE: 5.68, SSIM: 0.99 (d) OBNLM MSE: 12.9, SSIM: 0.98.

This experiment highlights an important point - evaluation methods can only account for certain specific features of an image and may not accurately assess image quality. In the results, MSE would suggest SRAD is superior, whilst SSIM and USDSAI suggest SBF. It is also apparent that this contradicts the work by [8] - most likely due to the subjectivity of both denoising and evaluation methodologies and differences between data sets. Arguably, SBF, SRAD and OBNLM are all similarly effective, adjustments to their parameters (iterations, smoothing, etc.) can be made repeatedly in order to improve one metric such as USDSAI, but this can lead to reduction in SSIM.

Method	MSE		SSIM		USDSAI	
	Best	Avg.	Best	Avg.	Best	Avg.
SBF	3.47	3.69	0.68	0.59	1.94	1.52
SRAD	2.77	2.82	0.81	0.75	2.01	1.72
OBNLM	2.83	3.10	0.69	0.43	1.42	1.15
Median	3.11	3.15	0.60	0.46	1.23	1.13

Table 2: Evaluation of various denoising algorithms.

5 Discussion

Noise is a prevalent issue in ultrasound, however there are a number of effective state-of-the-art methodologies for denoising, as well as techniques for evaluating them. Anisotropic and NL-means based algorithms are the most effective and recently developed methods for denoising in ultrasound are better suited to speckle compared to traditional approaches.

The term 'noise' for ultrasound often implies speckle exclusively. However, contrary to the typical desire to remove noise, statistical analysis of speckle has the potential for classifying regions correlating to anatomical structure. This could be used to segment regions or identify potential seed points for further processing. Raeth [24] observed that for ultrasound, computer analysis was superior to human observers and whilst speckle may seem to degrade an ultrasound image visually, it may provide useful information.

5.1 Inference from Speckle

Typically treated as noise, speckle might be a source of data and research has been conducted into characteristics of speckle distribution as a means for tissue differentiation [28] as it is the deterministic behaviour of waves in a particular environment. In a coherent system (such as ultrasound) images are formed through constructive and destructive interference of waves which results in fluctuation of amplitude (characterised in ultrasound by change in brightness). Ultrasound waves are scattered by surfaces or features in the field of view, the interference of these scattered waves manifests itself as speckle.

With appropriate analysis, speckle patterns may provide a means of inferring structure and other anatomical information; this has been demonstrated for a number of applications. One major research area is speckle tracking [4, 20] such as the clinically applied Laser Speckle Contrast Analysis (LASCA), which monitors change in speckle over time to determine blood flow [25]. Marti et al. [19] present a technique based upon an ellipsoid discriminant function to classify patches and generate speckle probability images, showing a clear correlation of speckle with anatomical structure. Correlation between speckle and intramuscular fat was visually observed in cattle during the 1980s [6] and tissue classification may be achieved with a number of methods such as wavelet-based filters, which have been applied to prostate images [10]. These analysis techniques typically utilise *speckle extraction* algorithms; these aim to separate an ultrasound mage into diffuse and coherent components. One such example is the Wold decomposition that thresholds an ultrasound signal [11] that has been applied to breast images for classification of normal and diseased tissue [11].

With this information in mind, it would be prudent for any researcher attempting to reduce the appearance of speckle to consider the application of speckle classification and probability estimation as a part of their processing pipeline.

6 Conclusions

We have reviewed a variety of major topics pertaining to ultrasound noise; specifically denoising but but we have also introduced the importance of speckle. Density maps would be useful to perfect as a means of selecting seed points for boundary delineating algorithms as well as providing visual clues to clinicians in real time. Work in ultrasound elastography has typically used speckle tracking [21, 26] whilst an entirely separate piece of research used variation in speckle to allow adaptive processing [1], increasing efficiency and reducing the need for unnecessary denoising in highly speckled areas. Deriving information from speckle is challenging, but could yield useful results in a variety of unexpected applications.

In conclusion, there are clearly superior two state-of-the-art algorithms (SRAD and OBNLM) for denoising though it is difficult to assess their efficacy. Before we can perform a truly conclusive evaluation, new metrics for assessing image enhancement must be developed as inconsistencies in the techniques of SSIM, MSE, etc demonstrate they cannot be relied upon. However, it seems that non-local techniques and anisotropic-diffusion are key areas for future work relating to denoising. A combinatory approach may be most appropriate, and indeed, non-local anisotropic diffusion has been applied somewhat recently by Yu [31] to restore conventional images whilst Krissian and Aja-Fernandez [16] have performed extended SRAD to incorporate local statistics.

Acknowledgements

We thank The Prostate Cancer Charity in association with the Hoover Foundation for funding this project.

References

[1] K.Z. Abd-Elmoniem, A.-B.M. Youssef, and Y.M. Kadah. Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion. *Biomedical Engineering, IEEE Transactions on*, 49(9):997 –1014, September 2002.

[2] A. Achim, A. Bezerianos, and P. Tsakalides. Novel Bayesian multiscale method for speckle removal in medical ultrasound images. *IEEE Transactions on Medical Imaging*, 20:772–783, August 2001.

[3] Charles C. Beckner and Charles L. Matson. Using mean-squared error to assess visual image quality. In *Optics & Photonics*, volume 6313. The International Society for Optical Engineering., August 2006.

[4] L. N. Bohs and G. E. Trahey. A novel method for angle independent ultrasonic imaging of blood flow and tissue motion. *IEEE Transactions on Biomedical Engineering*, 38: 280–286, March 1991.

[5] C. Boncelet. *Handbook of Image and Video Processing (Communications, Networking and Multimedia)*. Academic Press, Inc., Orlando, FL, USA, 2005.

[6] J.R. Brethour. Relationship of ultrasound speckle to marbling score in cattle. *Journal of Animal Science*, 68:2603–2613, 1989.

[7] A. Buades, B. Coll, and J.M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Model Simulations*, 4:490–530, 2005.

[8] P. Coupé, P. Hellier, C. Kervrann, and C. Barillot. Nonlocal means-based speckle filtering for ultrasound images. *IEEE Transactions on Image Processing*, 18:2221–2229, October 2009.

[9] F.P.X. de Fontes, G. Andrade Barroso, Pierrick, and P. Hellier. Real time ultrasound image denoising. *Real-Time Image Processing*, 6:15–22, March 2011.

[10] L. de Marchi, N. Testoni, and N. Speciale. Prostate tissue characterization via ultrasound speckle statistics. In *IEEE International Symposium on Signal Processing and Information Technology*, page 208 211, August 2006.

[11] G. Georgiou, F.S. Cohen, C.W. Piccoli, F. Forsberg, and B.B. Goldberg. Tissue characterization using the continuous wavelet transform. II: Application on breast rf data. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 48:364 – 373, March 2001.

[12] A. Horé and D. Ziou. Image quality metrics: PSNR vs. SSIM. In *20th International Conference on Pattern Recognition (ICPR)*, pages 2366–2369, August 2010.

[13] Q. Huynh-Thu and M. Ghanbari. Scope of validity of PSNR in image/video quality assessment. *Electronics Letters*, 44, 19 2008.

[14] Jørgen A. Jensen. FIELD: A Program for simulating ultrasound systems. In *10th Nordic-Baltic Conference on Biomedical Imaging*, volume 4, pages 351–353, 1996.

[15] L. Jong-Sen. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2:165–168, March 1980.

[16] K. Krissian and S. Aja-Fernandez. Noise-driven anisotropic diffusion filtering of mri. *Image Processing, IEEE Transactions on*, 18(10):2265 –2274, October 2009.

[17] D. Kuan, A. Sawchuk, T. Strand, and P. Chavel. Adaptive restoration of images with speckle. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35:373–383, March 1987.

[18] T. Loupas, W.N. McDicken, and P.L. Allan. An adaptive weighted median filter for speckle suppression in medical ultrasonic images. *IEEE Transactions on Circuits and Systems*, 36:129–135, January 1989.

[19] R. Martí, J. Martí, J. Freixenet, R. Zwiggelaar, J. C. Vilanova, and J. Barceló. Optimally discriminant moments for speckle detection in real B-scan images. *Ultrasonics*, 48: 169–181, July 2008.

[20] J. Meunier and M. Bertrand. Ultrasonic texture motion analysis: theory and simulation. *IEEE Transactions on Medical Imaging*, 14:293 –300, June 1995.

[21] J Ophir, B Cespedes, B Garra, H Ponnekanti, Y Huang, and N Maklad. Elastography: Ultrasonic imaging of tissue strain and elastic modulus in vivo. *European Journal of Ultrasound*, 3(1):49 – 70, 1996.

[22] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[23] C. Perreault and M. Auclair-Fortier. Speckle simulation based on B-mode echographic image acquisition model. In *Proceedings of the Fourth Canadian Conference on Computer and Robot Vision*, pages 379–386, 2007.

[24] U. Raeth, D. Schlaps, B. Limberg, I. Zuna, A. Lorenz, G. Van Kaick, W.J Lorenz, and B. Kommerell. Diagnostic accuracy of computerized B-scan texture analysis and conventional ultrasonography in diffuse parenchymal and malignant liver disease. *Journal of Clinical Ultrasound*, 13:87–99, 1985.

[25] G.J. Richards and J.D. Briers. Laser speckle contrast analysis (LASCA): a technique for measuring capillary blood flow using the first order statistics of laser speckle patterns. In *Biomedical Applications of Photonics (Digest No. 1997/124), IEE Colloquium on*, pages 11/1 –11/6, April 1997.

[26] A.R. Skovoroda, S.Y. Emelianov, M.A. Lubinski, A.P. Sarvazyan, and M. O'Donnell. Theoretical analysis and verification of ultrasound displacement and strain imaging. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 41(3):302 –313, May 1994.

[27] P.C. Tay, S.T. Acton, and J.A. Hossack. A stochastic approach to ultrasound despeckling. In *3rd IEEE International Symposium on Biomedical Imaging*, pages 221 –224, April 2006.

[28] R.F. Wagner, S.W. Smith, J.M. Sandrik, and H. Lopez. Statistics of speckle in ultrasound B-scans. *IEEE Transactions on Sonics and Ultrasonics*, 30:156–163, May 1983.

[29] Z. Wang and A. C. Bovik. Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26:98–117, January 2009.

[30] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.

[31] J Yu. Image de-noising based on nonlocal diffusion tensor. In *Information Assurance and Security, 2009. IAS '09. Fifth International Conference on*, volume 2, pages 501 –504, August 2009.

[32] Y. Yu and S. Acton. Speckle reducing anisotropic diffusion. *IEEE Transactions on image processing*, 11:1260–1270, 2002.

[33] Y. Yue, M.M. Croitoru, A. Bidani, J.B. Zwischenberger, and Jr. Clark, J.W. Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images. *IEEE Transactions on Medical Imaging*, 25:297–311, March 2006.

Multi-Scale Local Phase Features for Anatomical Object Detection in Fetal Ultrasound Images

Bahbibi Rahmatullah[1]
bahbibi.rahmatullah@eng.ox.ac.uk
Aris Papageorghiou[2]
aris.papageorghiou@obs-gyn.ox.ac.uk
J. Alison Noble[1]
alison.noble@eng.ox.ac.uk

[1] Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford, UK

[2] Nuffield Department of Obstetrics & Gynaecology, John Radcliffe Hospital, University of Oxford, UK

Abstract

The presence and absence of anatomical landmarks in fetal ultrasound image are commonly used as indicators in finding the optimal ultrasound plane for fetal biometric measurement. We propose a machine learning framework for the automated detection of two important anatomical landmarks (stomach and umbilical vein) in fetal abdominal ultrasound images. Our proposed approach combines the unary features extracted from the multi-scale local phase image with the Haar features acquired from the intensity image. The results presented indicate an improvement over prior intensity-based approach.

1 Introduction

Comprehensive ultrasound examination during pregnancy is required for the clinical purpose of dating pregnancy, monitoring the growth of the fetus and detecting any abnormalities. During the scan, the sonographer scans the thigh, abdomen and head area to get the standard fetal biometric measurements (e.g. the biparietal diameter (BPD), occipito-frontal diameter (OFD), head circumference (HC), abdominal circumference (AC), and femur length (FL)). These biometric measurements are then plotted against the pre-determined growth chart in order to check the normality in the fetal growth. It is very important to have accurate measurements in order to avoid the risks of a false negative (failure to detect abnormality in growth) or false positive which might lead to medical intervention and unnecessary maternal anxiety.

The accuracy of the measurements depend on the correct positioning of the on-screen measurement calliper in the image acquired from a standard plane. The optimality of the acquired image can be checked by identifying the visibility of certain anatomical structures inside the fetus. For example, in a standard fetal abdominal ultrasound scan, the stomach bubble (SB) is visible and the umbilical vein (UV) is at the one-third position from the abdomen wall [1].

The work presented in this paper investigates the effect of introducing a new feature sets into the machine learning framework for the localization of the two important anatomical landmarks (stomach and umbilical vein) in fetal abdominal ultrasound images.

Previous work considered only features from intensity images [2]. Our approach uses the combination of features extracted from multi-scale local phase image and Haar features acquired from intensity images in the training of the object detector using an adaptive boosting (AdaBoost) learning algorithm.

2 Local Phase Features

In ultrasound image analysis, the feature information extracted from local phase-based processing of images has proven beneficial for a variety of image analysis task [3, 4]. Local phase is invariant to changes in image brightness or contrast, making it particularly suitable for ultrasound images, known for its relatively low signal-to-noise ratio, artefacts and shadowing. These works show that a local-phase based method outperforms the conventional intensity-based methods for feature detection in ultrasound images.

The local phase (φ) is defined as:

$$\varphi(x) = \arg\big(f_A(x)\big) = \arctan\left(\frac{Im(f_A(x))}{Re\big(f_A(x)\big)}\right) \tag{2}$$

where $f_A(x)$is the complex analytic signal, $Re[\cdot]$ and $Im[\cdot]$ correspond to the real and complex portion of the signal, respectively.

In this implementation, we used an approach known as the monogenic signal [5]. The monogenic signal is an isotropic extension of the analytic signal which preserves the core properties of the 1-D analytic signal that decomposes a signal into information about its structure (local phase) and energy (local amplitude). The analytic signal is generated using an isotropic vector valued odd filter known as the Riesz transform, which is a generalization of the Hilbert transform for higher dimensional signals. The spatial representations of these filters are as follow:

$$h_1(x,y) = \frac{-x}{2\pi(x^2+y^2)^{3/2}} \tag{3}$$

$$h_2(x,y) = \frac{-y}{2\pi(x^2+y^2)^{3/2}} \tag{4}$$

In practice, the image $I(x,y)$ is first convolved with an even isotropic band-pass filter $b(x,y)$ that produces the even component of the monogenic signal:

$$even(x,y) = I_b(x,y) = b(x,y) * I(x,y) \tag{5}$$

The bandpassed image $I_b(x,y)$ is then filtered with the Riesz filter to produce the odd components:

$$odd_1(x,y) = h_1(x,y) * I_b(x,y)$$

$$odd_2(x,y) = h_2(x,y) * I_b(x,y)$$

$$odd(x,y) = \sqrt{\big(odd_1(x,y)\big)^2 + \big(odd_2(x,y)\big)^2} \tag{6}$$

The local phase $\varphi(x,y)$ of the image $I(x,y)$ is produced through the following definition:

$$\varphi(x,y) = arctan\left(\frac{even(x,y)}{odd(x,y)}\right) \tag{7}$$

The multi-scale local phase image $\varphi_{MS}(x,y)$ is formed by averaging the local phase computed at different scales:

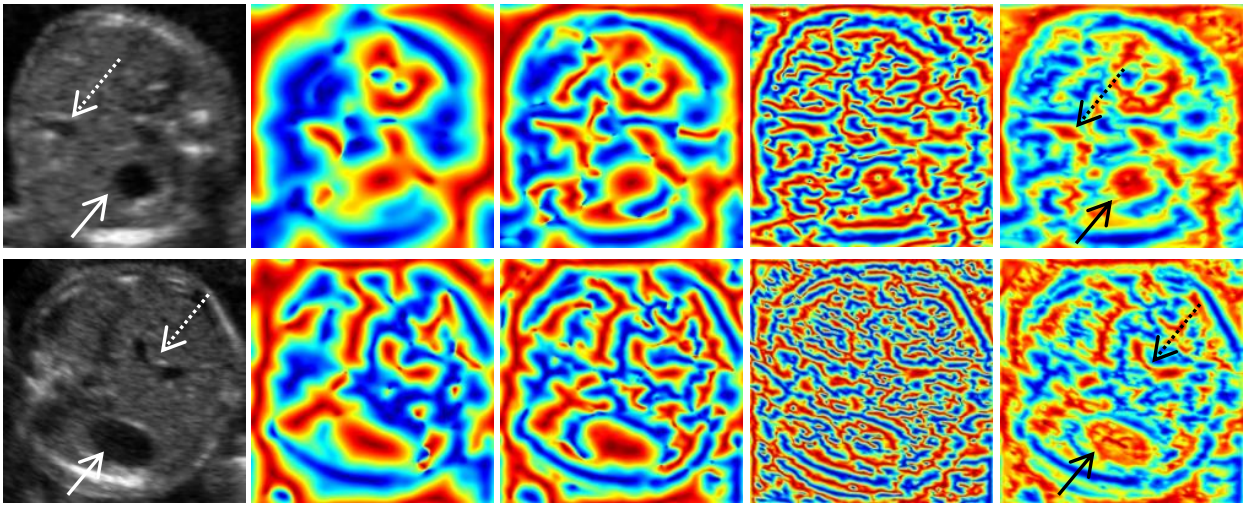
$$\varphi_{MS}(x,y) = \frac{1}{N}\sum_{s} \varphi_s(x,y) \tag{8}$$

where s represents the scales and N is the total number of scales.

For the band-pass filter, we used a log-Gabor filter because it allows arbitrarily large bandwidth zero DC filters to be constructed. In 1-D, the log-Gabor filter has a transfer function of the form:

$$G(\omega) = exp\left(-\frac{log^2(\omega/k)}{2log^2(\sigma_\omega)}\right)$$

where k is the centre frequency of the filter (which is inversely related to the scale of the filter), and $0 < \sigma_\omega < 1$ is related to the spread of the frequency spectrum in a logarithmic function. The following parameters produced the best empirical results: σ_ω= 0.50 and 3 scales of filter wavelength: [250 150 50] pixels. Figure 1 shows the effect of using single-scale filter and multi-scale filter in producing the local phase (LP) image.



(a) original (b) scale = 250 (c) scale = 150 (d) scale = 50 (e) multi-scale

Figure 1: Example of local phase images produced by using different scales of filter

The stomach (*solid arrow*) and the umbilical vein (*dashed arrow*) are shown to have a better identification and separation from the surroundings in (e) the multi-scale local phase image as compared to (b-d) other single-scale local phase images.

3 Haar features

Features derived from Haar wavelets have been commonly used for representing the information in the image region. There are two motivations for using Haar features instead of pixel intensities directly. Firstly, features encode domain knowledge better than pixels for different types of visual patterns including objects in ultrasound images. The other reason is that a feature-based system can be much faster than a pixel-based system due to the use of the integral image. In their framework of object detection, Papageorgiou et al. [6] proposed the Haar-like features as a basis for image representation which was later

employed in a face detection technique [7]. The set of rectangle features provide a rich image representation which enables an effective learning process.

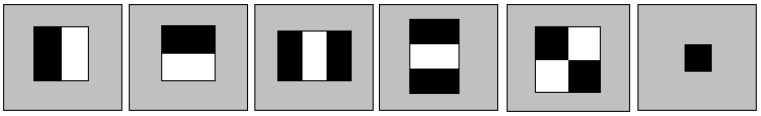


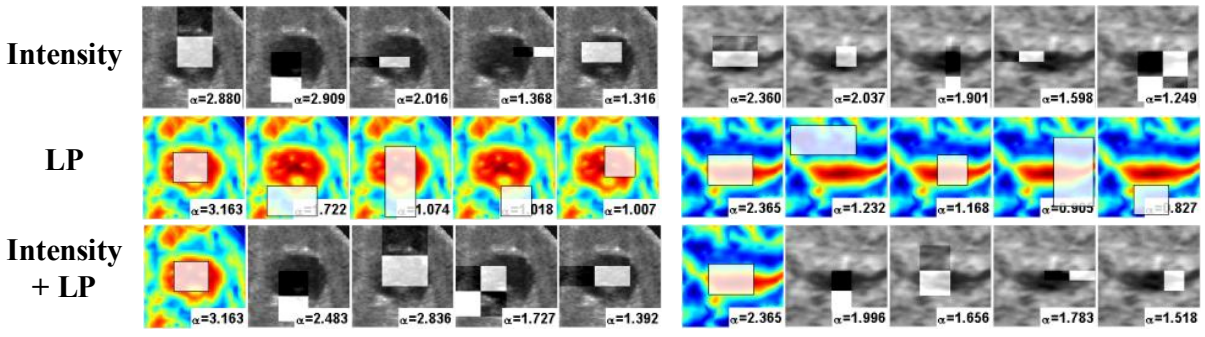
Figure 2: The unary, two-, three- and four- rectangles prototypes used in our implementation. The features are calculated by summing the value in black regions and subtracted with the totals in white regions (except for unary feature).

4 AdaBoost Learning Algorithm

For feature selections and classifier training, we used the AdaBoost [8] algorithm which is an established and proven effective method in object detection problems [7]. AdaBoost forms a strong hypothesis through linear combination of weak classifiers that are derived from the supplied pool of extracted features. The algorithm only requires the number of iterations (boosting) and automatically selects the most representative features during the training process.

The positive training samples for the classifier were cropped from the image regions that contain the anatomical object. Regions cropped from the background and the negative images (that do not contain the anatomical object) were used as negative training samples. We obtained 633 positive samples and 1032 negative samples for the stomach, and 448 positive samples and 965 negative samples for the umbilical vein training set.

Two different feature sets (*Intensity features* and *Local phase (LP) features*) were then extracted from the training dataset. Boosting process was performed separately on each feature set and also on the combination of both feature sets (*Intensity+LP*). The first five features chosen by the boosting algorithm are shown in Figure 3. Note that in both the stomach and the umbilical vein trained models, the local phase features were the first features selected by the learning algorithm and assigned with large classifier weights (α) which indicate high discriminating power of the features.



(a) (b)

Figure 3: The first five features selected by AdaBoost for (a) the stomach and (b) the umbilical vein detection superimposed on the sample images. The greyscale and the coloured images indicate that the features selected are from the intensity feature set and local phase feature set, respectively.

For testing, exhaustive scan using the sliding window technique was performed (for all possible translations and a sparse set of scales) to find the anatomical object in the test image. The trained classifier was applied to all sub-windows within an image and takes the maximum of the classification score (α_T) as indication of the presence or absence of an object.

5 Data Acquisition

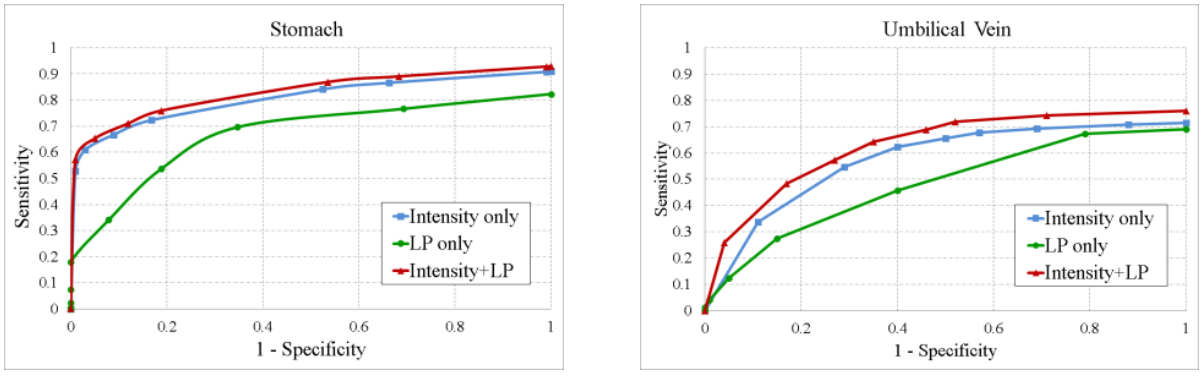
The fetal abdominal images used in this work were randomly selected from an ongoing clinical study database for fetal growth. The study involves normal singleton pregnancy from over 4000 healthy pregnant women screened at study entry to exclude risk factors for intrauterine growth restriction and overgrowth. Philips HD9 ultrasound machine with a 2-5MHz 2D probe are used by ultrasonographers trained to follow standardized procedures from study. The testing datases consisting of 2384 images were labeled after consultation with trained sonographers.

6 Results

The performance of the detection methods trained with the three different feature sets were compared using ROC curves, as shown in Figure 4. Area under the curve (AUC) and balanced accuracy [$(specificity + sensitivity)/2$] were summarized in Table 1.

Table 1: The performance of the detection methods using three different feature sets.

	Area under the curve (AUC)		Balanced Accuracy (%)	
Feature sets	Stomach	Umbilical Vein	Stomach	Umbilical Vein
Intensity	0.80	0.57	78.94	62.80
LP	0.66	0.48	67.40	56.18
Intensity + LP	0.83	0.63	80.14	65.16



(a) (b)

Figure 4: ROC plots for the detection of (a) the stomach and (b) the umbilical vein.

In comparison to the intensity-based features method, the proposed approach (*Intensity + LP* feature sets) achieved an increase of 1.20% and 2.36% in the accuracy of the stomach and the umbilical vein detection, respectively. These percentages translate to an increase of 100 true positive (TP) detections of stomach but with a decrease of 2 true

negative (TN) cases. For umbilical vein, it accounts for an increase of 62 TP and 2 TN detections.

7 Conclusions

In this work we have introduced a feature sets derived from multi-scale local phase images and intensity images into the machine learning framework for detecting anatomical landmarks (stomach and umbilical vein) in fetal abdominal ultrasound images. Quantitative results showed improved performance in comparison to the existing approach. This enhancement will in turn enable a more accurate quality assessment of fetal biometric ultrasound scans for improving the accuracy and the reproducibility of the fetal biometric measurements. This work is being extended to utilize other information available from local phase values (such as feature symmetry and asymmetry) and also the detection of objects in other fetal biometry scan areas.

References

[1] L. S. Chitty, D. G. Altman, A. Henderson, and S. Campbell, "Charts of fetal size: 3. Abdominal measurements," *British Journal of Obstetrics and Gynaecology,* vol. 101, pp. 125-131, 1994.

[2] B. Rahmatullah, I. Sarris, A. Papageorghiou, and J. A. Noble, "Quality control of fetal ultrasound images: Detection of abdomen anatomical landmarks using AdaBoost," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Chicago, IL, 2011, pp. 6-9.

[3] V. Grau and J. A. Noble, "Adaptive multiscale ultrasound compounding using phase information," vol. 3749 LNCS, ed. Palm Springs, CA, 2005, pp. 589-596.

[4] M. Mulet-Parada and J. A. Noble, "2D+T acoustic boundary detection in echocardiography," in *Medical Image Analysis*, 2000, pp. 21-30.

[5] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transactions on Signal Processing*, vol. 49, pp. 3136-3144, 2001.

[6] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, "Pedestrian detection using wavelet templates," in *Computer Vision and Pattern Recognition (CVPR)*, 1997, pp. 193-199.

[7] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision,* vol. 57, pp. 137-154, 2004.

[8] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Journal of Computer and System Sciences,* vol. 55, pp. 119-139, 1997.

Segmentation and Mesh Generation for Patient-Specific Flow Modelling in Carotid Arteries

Igor Sazonov[1]
i.sazonov@swansea.ac.uk
Xianghua Xie[2]
x.xie@swansea.ac.uk
Si Yong Yeo[1]
genyeoius@gmail.com
Rhodri Bevan[1]
r.bevan@swansea.ac.uk
Raoul van Loon[1]
r.vanloon@swansea.ac.uk
Perumal Nithiarasu[1]
p.nithiarasu@swansea.ac.uk

[1] College of Engineering
Swansea University
Swansea, SA2 8PP, UK

[2] Department of Computer Science
Swansea University
Swansea, SA2 8PP, UK

Abstract

Segmentation and mesh generation are two important stages in patient-specific computational modelling. In this paper, we present our approach to these two problems in modelling blood flow haemodynamics in human carotid arteries in order to improve the understanding of the impact of flow dynamics to carotid disease formation and progression. The segmentation of carotid arteries in CT scans is carried out using a deformable model based on a geometrical potential force that is derived from generalising image gradient vector interactions across the image domain. It is shown to be robust towards image noise interference, weak edge, and arbitrary initialisations. The mesh generation involves surface meshing and volume meshing. The surface meshes are computed from the implicit function obtained from the deformable modelling. The near-wall volumetric mesh is required in order to resolve the viscous boundary layer in flow studies. A flow study on severe stenosis is provided which involves segmenting the carotid geometries, generating valid surface and volume meshes, and a flow solver.

1 Introduction

It is known that vascular diseases such as stenosis and aneurysms are often associated with changes in blood flow patterns and the distribution of wall shear stress (WSS). It is however extremely complicated to measure those flow parameters *in vivo* at present time. Numerical modelling of the haemodynamics, based on the geometries obtained from 3D scans, can provide valuable insights which may help to development of efficient treatment methods. In

recent years, computational fluid dynamics (CFD) has been increasingly used for patient-specific modelling of blood flow in vascular structures, but so far there has been very limited applications of computational haemodynamics in clinical practice. This is largely due to the challenges involved in the design of an integrated framework which can robustly and accurately automate the computational modelling process, which includes image segmentation, mesh generation, and computational fluid dynamics simulation.

Segmentation and subsequent mesh generation have shown to be an intricate procedure and often application dependent. Due to lack of generic methods and robust automation, performing computational modelling on a large scale of data is particularly difficult. In this work, we adopt a generic segmentation algorithm to obtain the carotid geometries from CT scans and a mesh generation approach which attempts to minimise user interference. We demonstrate the proposed approach on analysing wall shear stress pattern in severe stenosis. The results from this preliminary study show that the proposed method is a promising approach to an integrated, generic and automated patient-specific computational modelling.

2 Image segmentation

The GPF method. Among many others, deformable modelling is a popular approach to image segmentation, e.g. [6, 10]. Conventional techniques suffer from weak edge, image noise and convergence issues. For instance, in [6] a constant pressure force is necessary in order to improve its capture range, resulting in monotonic expanding or shrinking of the mode that is problematic. In this work, we adopt the Geometric Potential Force (GPF) based deformable model [12] to carry out the 3D segmentation. It is a purely data driven method which is generic but has shown to be robust towards weak edge and image noise interference, and has the ability to cope with difficult initialisations, such as cross boundary initialisation and converging to narrow and long passages which is common in vascular structures. To cope with the size of the dataset, we use an efficient implementation to speed up the segmentation.

Let the 3D grey level image be described by function $I(\mathbf{x})$ where $\mathbf{x} = [x,y,z]^T \in \mathscr{D}$ is a point in the image domain $\mathscr{D}$. Let Ω be an object to be segmented. We employ the level set method in which the object boundary $\partial\Omega(t)$ is defined through level set function $\Phi(\mathbf{x},t)$: $\partial\Omega(t) = \{\mathbf{x} : \Phi(\mathbf{x},t) = 0\}$. The most commonly used PDE to compute $\Phi(\mathbf{x},t)$ can be formulated as:

$$\frac{\partial\Phi}{\partial t} = \alpha\, g(\mathbf{x})\, \kappa(\mathbf{x},t)\, \|\nabla\Phi\| - (1-\alpha)(\mathbf{F}(\mathbf{x})\cdot\nabla\Phi) \tag{1}$$

where α is a real constant, $g(\mathbf{x}) = 1/(1+\|\nabla I\|)$ is the edge stopping function, $\kappa(\mathbf{x},t) = \nabla\cdot\hat{\mathbf{n}}$ is the mean curvature of surface $\Phi = const$, $\hat{\mathbf{n}} = \nabla\Phi/\|\nabla\Phi\|$ is the unit normal vector to that surface, $\mathbf{F}(\mathbf{x}) = [F_x, F_y, F_z]^T$ is the flow function determined by image I. The geometrical potential force (GPF) proposed in [12] is computed as:

$$\mathbf{F}(\mathbf{x}) = G(\mathbf{x})\cdot\hat{\mathbf{n}}, \qquad G(\mathbf{x}) = P.V.\int_{\mathbf{x}'\in\mathscr{D}} \nabla I(\mathbf{x}')\cdot\frac{\mathbf{x}-\mathbf{x}'}{\|\mathbf{x}-\mathbf{x}'\|^{n+1}}\,\mathrm{d}^n\mathbf{x}', \tag{2}$$

where G is the scalar potential and *P.V.* denotes the principle value (as the integral can diverge in vicinity $\mathbf{x}' = \mathbf{x}$), and $n = 3$ is the image dimension. This potential can be also represented as a convolution operation: $G = \nabla I * \mathbf{K}$ where $\mathbf{K} = P.V.\left(\mathbf{x}/\|\mathbf{x}\|^{n+1}\right)$ is the vector kernel (note that the convolution is united with the dot-product as ∇I and $\mathbf{K}$ are vectors). A discrete

analogue of the convolution kernel is $\mathbf{K}(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|^{n+1}$ if $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{0}$ if $\mathbf{x} = \mathbf{0}$. An efficient way to implement the convolution (2) is to apply the Fast Fourier Transform (FFT) method.

Equation (1) is solved numerically by using the finite difference method: using upwind scheme for the flow $\mathbf{F} \cdot \nabla\Phi$ and central differences for other derivatives. Advancing in time is performed by the forward Euler method. It is possible to show that in the 2D case with $n = 2$ this force coincides with a magnetic-active force [11]. Compared with the other image gradient based segmentation methods, this method is more stable to topological complexity of the object, to image noise, to model initialisation, to leakaged at fuzzy edges as shown in [12].

Improvements to the GPF method. When applied to 3D and large dataset it requires substantial amount of memory and computation still can be demanding even using FFT, as it is necessary to store arrays for all components of ∇I, for all components of kernel $\mathbf{K}$, and twice more for the results of the FFT evaluation. This drawback was an essential restriction of the method proposed in [12] if applied to the typical size of 3D medical scans.

To alleviate this, we re-write the integrand in (2) so that the integral is represented as convolution of scalar functions. Let $\tilde{I}(\mathbf{k})$ denote the Fourier transform of the image $I(\mathbf{x})$:

$$\tilde{I}(\mathbf{k}) = \int I(\mathbf{x}) e^{i\mathbf{k}\mathbf{x}} \mathrm{d}^n\mathbf{x} \tag{3}$$

where $\mathbf{k} = [k_x, k_y, k_z]^T$ (if $n = 3$). Then $\tilde{\nabla} I = i\mathbf{k}\tilde{I}$ and $\tilde{G} = (i\mathbf{k}\tilde{I}) \cdot \tilde{\mathbf{K}}^*$. The last equation can be re-written as

$$\tilde{G} = \tilde{I}(-i\mathbf{k} \cdot \tilde{\mathbf{K}})^* \qquad \Longleftrightarrow \qquad G = -I * (\operatorname{div} \mathbf{K}) \tag{4}$$

Scalar kernel $K = (\operatorname{div} \mathbf{K})$ behaves as $\|\mathbf{x}\|^{-(n+1)}$ when $\mathbf{x} \neq \mathbf{0}$ and has a strong singularity at $\mathbf{x} \neq \mathbf{0}$ such that $\int K(\mathbf{x})\, \mathrm{d}^n\mathbf{x} = 0$. In virtue of these properties, the discrete analogue is

$$K(\mathbf{x}) = \begin{cases} \|\mathbf{x}\|^{-(n+1)}, & \mathbf{x} \neq \mathbf{0} \\ -S, & \mathbf{x} = \mathbf{0} \end{cases}, \qquad S = \sum_{\mathbf{x} \in \Omega, \mathbf{x} \neq \mathbf{0}} \|\mathbf{x}\|^{-(n+1)}. \tag{5}$$

The potential G is a linear transform of I. It can be precomputed before evolving the deformable model (2). An example of 3D segmentation using this method is shown in Figure 1. It producd visually the same result as the original GPF method but it is more efficient and uses three times less memory.

3 Mesh generation

Despite a significant effort in developing robust patient-specific meshing methods, generation of a valid mesh automatically and rapidly is a challenge. A valid mesh for cardiovascular flows should be sufficiently fine to capture WSS with minimal error associated with point distribution. Using a coarse and purely unstructured mesh can produce a WSS distribution that is far from a converged solution. Although mesh convergence is not always straightforward to carry out, designing a mesh by taking into account all the necessary factors, including boundary layer, is essential to obtain results with high level of accuracy.

In patient-specific modelling, it is essential to have smooth transfer of image segmentation results to the meshing stage. In the majority of works reported, this link is often not well defined. Patient-specific meshing is a growing area of research, and currently, there is no universal way of satisfying all the meshing requirements via a single algorithm. In particulalr,

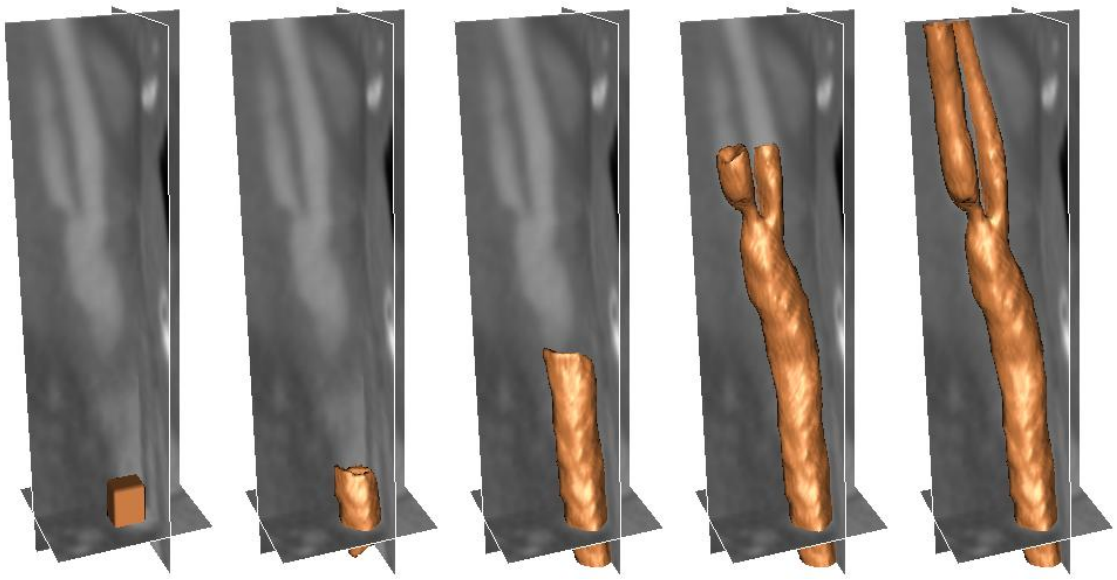


Figure 1: Segmentation of carotid artery from CT scan using the GPF deformable model with improved efficiency.

the interface between image processing and a valid surface generationis not satisfactorily addressed.

In standard engineering applications, the object boundary is rigorously defined and described analytically or piecewise analytically. A well-defined boundary allows to simplify the construction of a surface mesh and the corresponding cosmetics. In patient-specific medical objects, the surface is not well-defined and it may not be easily described analytically or piecewise analytically without compromising important surface features. Hence, alternative approaches have to be used and often it is essential to approximate the surface of a medical object. In this work, we generate surface meshes using the converged level set function from the segmentation, followed by boundary layer meshing.

The segmented domain is to be discretised, i.e. to build a conformal mesh of polyhedron elements approximating the domain as accurate as possible. We use a tetrahedron mesh with triangular surface elements in our simulation. The performance of a solution scheme depends significantly on the mesh quality. The elements should be of appropriate size, quality and shape to obtain a sensible and accurate numerical solution. The meshing stage also includes the accurate determination of the object inlet and outlets which should be orthogonal to the blood-vessel.

Surface meshing. The surface mesh generation procedure has the following stages.
1) The level set function calculated by the segmentation is used to determine the approximate axis of the arteries and their branches. These skeleton-like structures are useful in automatic identification of the geometrical features of the vessel, such as stenosis, aneurysms, bifurcations and kinks. It is also helpful for determining the appropriate orientations for the inlet and outlet surfaces. We apply voxel thinning algorithm to perform the skeletonisation, followed by axis smoothing to remove minor irregularities from skeletonisation.
2) Representation of the object surface through a continuous level set function gives an opportunity to employ an advanced Marching Cube (MC) algorithm [2] to obtain the surface. This gives a much smoother surface than the standard MC method.
3) Once a satisfactory surface representation is established and the geometrical skeleton

is generated, we are then bale to obtain the valid inlet and exit boundary surfaces. Often clipping is carried out manually, but in this work, we perform this automatically by using the generated surface and skeletons.

4) The mesh obtained by the advanced MC method often contains a large number of ill-shaped elements. In addition, the triangular element size can be smaller or larger than the element size required by the flow solver. Therefore, the initial mesh needs some further refinement, which creates a surface mesh for the same domain with much better element quality and at a prescribed element size. Here, we apply three procedures to improve the mesh quality:

(i) Mesh smoothing (no topological changes). We use the Taubin smoothing instead of traditional Laplace smoothing which can cause undesired shrinking of the objects.

(ii) Edge swapping. Topology based edge swapping is employed. The edge is swapped if the nodal index (number of contiguous modes) of vertices of triangles containing given edge becomes closer to an ideal value that is 6.

(iii) Splitting/contraction of edges. A too long edge is split by inserting a new node on a surface near the midpoint of the edge. A too short edge is reduced to a point on a surface near the midpoint of the contracted edge. Following the splitting and contraction the local mesh cosmetics is applied.

Volume meshing. To resolve the viscous boundary layer in flow studies, a special structure of the near-wall volumetric mesh is required. The elements built in the direction of inward normal to the wall should be essentially smaller than the element size in the inner volume of an artery. The short edge of a near-boundary tetrahedra should be directed normal to the wall. Such a quasi-structured, near-boundary mesh consists of N sub-layers with decreasing thickness towards the boundary. The mesh can be generated if we build triangular prisms at every face, by dividing every prism into N smaller prisms with height ratio f between two neighbouring prisms. To obtain a tetrahedron mesh, every triangular prism can be split into three tetrahedra. In [1], it is shown that $N = 10$ sub-layers is sufficient to evaluate WSS with the desired accuracy. Additional 3D smoothing technique based on 3D Lloyd's [3] iterations is also applied. Figure 2 gives an example of the meshing process and it also shows the computed WSS distribution.

4 Flow solution

The waveform of the inlet maximal velocity is extracted from ultrasound measurements [4], and the boundary conditions in the inlet are computed as a generalisation of the Womersley solution [9] onto non-circular pipe [7]. The boundary condition in the inlet is specified such that the mass flow through the considered part of the blood vessel is appropriate and the outlets have a realistic split of the mass flow at the branches. The vessel wall a considered a rigid in this study. In order to model biofluid flows, the locally conservative Galerkin (LCG) method is employed within the characteristic based split (CBS) scheme [5]. To account for possible turbulence, a Sparlart-Allmaras (SA) turbulence model is employed [8]. The influence of the boundary mesh parameters on convergence of the solution have been investigated. From the CFD simulation, information on the flow patterns, pressure distribution, and wall shear stress (WSS) has been determined, together with further WSS derived parameters which have been linked to intima-media thickness and risk of atheroma.

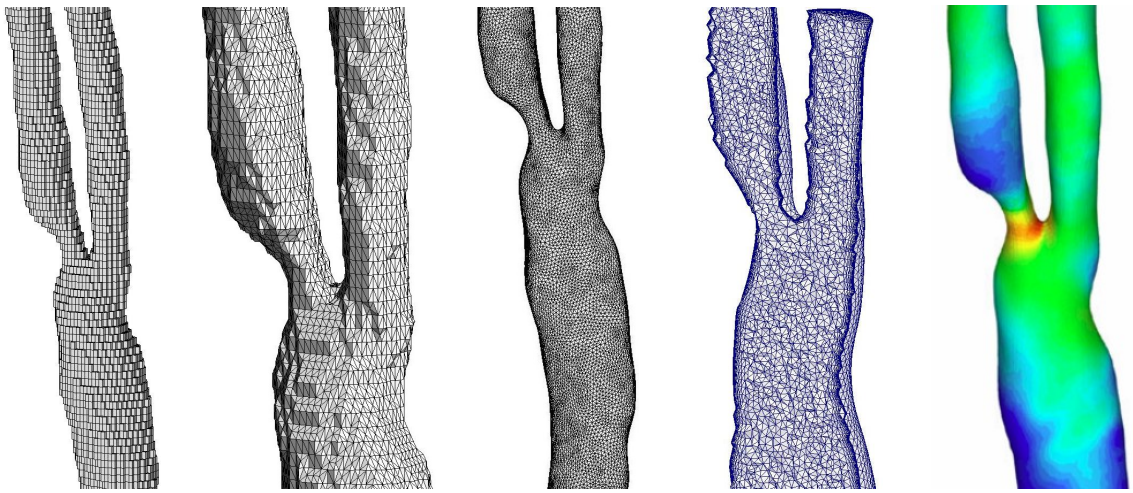


Figure 2: Meshing stages: (a) binary image, (b) initial mesh, (c) surface mesh after cosmetics, (d) a cut off view of the volumetric mesh. (e): time-averaged WSS distribution.

5 Conclusion

In this paper, we briefly presented a pipeline for patient-specific computational modelling from medical image scan, to segmentation, to mesh generation and finally to computational simulation. The focus of this work is on segmentation, which is based on an extension of the GPF model, and computational mesh generation. Automatic mesh generation is particularly difficult due to the complexity of the structure and high mesh quality demand from finite element based simulation. The proposed method showed promising results towards a fully automated segmentation and meshing process.

References

[1] R.L.T. Bevan, P. Nithiarasu, R. Van Loon, I. Sazonov, H. Luckraz, and A. Garnham. Application of a locally conservative Galerkin (LCG) method for modelling blood flow through a patient-specific carotid bifurcation. 64:1274–1295, 2010.

[2] E. V. Chernyaev. Marching cubes 33: Construction of topologically correct isosurfaces. Technical report Cern CN 95-17, CERN, 1995.

[3] Q. Du and D. Wang. Tetrahedral mesh generation and optimization based on CVT. *International Journal for Numerical Methods in Engineering*, 56:1355–1375, 2003.

[4] D.W. Holdsworth, C.J.D. Norley, R Frayne, DA Steinman, and BK Rutt. Characterization of common carotid artery blood-flow waveforms in normal human subjects. *Physiological Measurement*, 20(3):219–240, 1999.

[5] C.B. Liu and P. Nithiarasu. The explicit characteristic based split (cbs) method for viscoelastic flow past a circular cylinder. *International Journal for Numerical Methods in Fluids*, 57:157–176, 2008.

[6] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modelling with front propagation: A level set approach. *IEEE T-PAMI*, 17(2):158–175, 1995.

[7] I. Sazonov, S.-Y. Yeo, R. L. T. Bevan, X. Xie, R. van Loon, and P. Nithiarasu. Modelling pipeline for subject-specific arterial blood flow -A review. *International Journal for Numerical Methods in Biomedical Engineering*, 27(8):1167–1184, 2011.

[8] P.R. Spalart and S.R. Allmaras. A one-equation turbulence model for aerodynamic flows. *AIAA paper 92-0439*, 1992.

[9] J.R. Womersley. Oscillatory flow in arteries: the constrained elastic tube as a model of arterial flow and pulse transmission. *Physics in Medicine and Biology*, 2:178–187, 1957.

[10] X. Xie. Active contouring based on gradient vector interaction and constrained level set diffusion. *IEEE T-IP*, 19(1):154–164, 2010.

[11] X. Xie and M. Mirmehdi. MAC: Magnetostatic active contour model. *IEEE T-PAMI*, 30(4): 632–647, 2008.

[12] S. Y. Yeo, X. Xie, I. Sazonov, and P. Nithiarasu. Geometrically induced force interaction for three-dimensional deformable models. *IEEE T-IP*, 20(5):1373–1387, 2011.

On computation of diffusion and fibre orientation distribution functions in high angular resolution diffusion imaging

Bartosz P. Neuman[1]
bartosz.neuman@nottingham.ac.uk
Christopher Tench[2]
christopher.tench@nottingham.ac.uk
Li Bai[1]
bai.li@nottingham.ac.uk

[1] School of Computer Science
University of Nottingham
Nottingham, UK

[2] Academic Division of Clinical Neurology
University of Nottingham
Nottingham, UK

Abstract

Diffusion weighted MRI is a non-invasive image technique for obtaining information about the neural architecture of the brain. The measured diffusion signal is highly correlated with the direction of the white matter tracts. Based on the transformed signal (either in the form of orientation distribution function or fibre orientation density) it is possible to estimate the fibre orientations, provided the algorithm used can cope with the noise corrupted diffusion weighted image. This paper reviews the methods used to compute the diffusion and fibre orientation distribution functions.

1 Introduction

Diffusion weighted MRI (DW-MRI) is a non-invasive imaging technique that allows to measure the displacement (diffusion) of water molecules. Applied to the brain, it can be used to recreate the white matter tracts [3], study brain connectivity [26], and detect early changes in the cerebral tissue [4].

Despite intensive research in the DW-MRI, very few methods have been used in clinical applications. One of those that were successful though, was a high angular resolution diffusion imaging (HARDI) protocol [13, 22, 32]. It allows to measure the diffusion signal in a clinically feasible way (*e.g.* provides a good trade-off between the scan time and amount of acquired data). The number of acquired volumes varies on application, but is usually between 30 and 60. Among the methods utilising HARDI there are: generalized diffusion tensor imaging [22], persistent angular structure MRI [17], q-ball imaging (QBI) [31], fibre orientation estimated using continuous axially symmetric tensors [2], and diffusion orientation transform [23].

This paper reviews a QBI and methods related to the computation of diffusion orientation distribution function (ODF) and fibre orientation density (FOD).

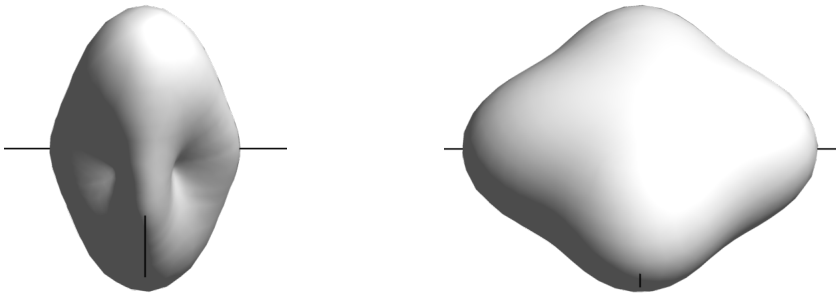


Figure 1: Relationship between the diffusion signal (left) and ODF (right), simulated 90° fibre crossing, 8th order SHS, $b = 3000$ mm^2/s.

2 Orientation distribution function

In diffusion MRI, the orientation distribution function (ODF) characterizes the 3D distribution of water diffusion, and is necessary to infer the fibre configuration. QBI [31, 33] is a model-independent reconstruction scheme for HARDI. The diffusion ODF Ψ is defined as the Funk-Radon transform of the diffusion signal E (see Figure 1 for visual relationship between E and Ψ) and can be computed as:

$$\Psi(\mathbf{u}) = \int_{\mathbf{q}\perp\mathbf{u}} E(\mathbf{q})dq, \tag{1}$$

where both $\mathbf{u}$ and $\mathbf{q}$ are unit directions.

The numerical approach to calculate the equator integral required more data than was available from HARDI (since the equator points did not coincide with the diffusion sampling points), and the diffusion signal was interpolated using the spherical radial basis function (sRBF) [12]. Subsequently, Anderson [2], Hess [15, 16], and Descoteaux [9] have independently and in parallel developed an analytical solution for the ODF reconstruction in QBI using spherical harmonic (SH) [21] basis function and Funk–Hecke theorem. First, the diffusion signal is approximated with a truncated spherical harmonic series (SHS) [1]:

$$\hat{E} = \sum_{l=0}^{n} \sum_{m=-l}^{l} c_{lm} Y_l^m(\theta,\phi), \tag{2}$$

with $\hat{E}$ being the approximation of signal E using nth order SHS, Y_l^m a spherical harmonic function and c_l^m a SH coefficient of order l and band m. The coefficients of the series are found using the spherical harmonic transform (SHT):

$$c_{lm} = \int_0^{2\pi} \int_0^{\pi} E(\theta,\phi) Y_l^m(\theta,\phi) \sin\theta d\theta d\phi. \tag{3}$$

In practice, the Equation 3 is rarely used and instead a discrete approximation of the exact solution is found using a linear least squares method:

$$\mathbf{c} = (\mathbf{Y}^{\mathrm{T}}\mathbf{Y})^{-1}\mathbf{Y}^{\mathrm{T}}\mathbf{E}. \tag{4}$$

Here $\mathbf{Y}$ is a SH design matrix, and $\mathbf{E}$ is a vector containing the measurements. The diffusion ODF can now be directly estimated from the SH representation of the diffusion signal as:

$$\Psi(\theta,\phi) \approx \sum_{l=0}^{n} \sum_{m=-l}^{l} 2\pi P_l(0) c_{lm} Y_l^m(\theta,\phi), \tag{5}$$

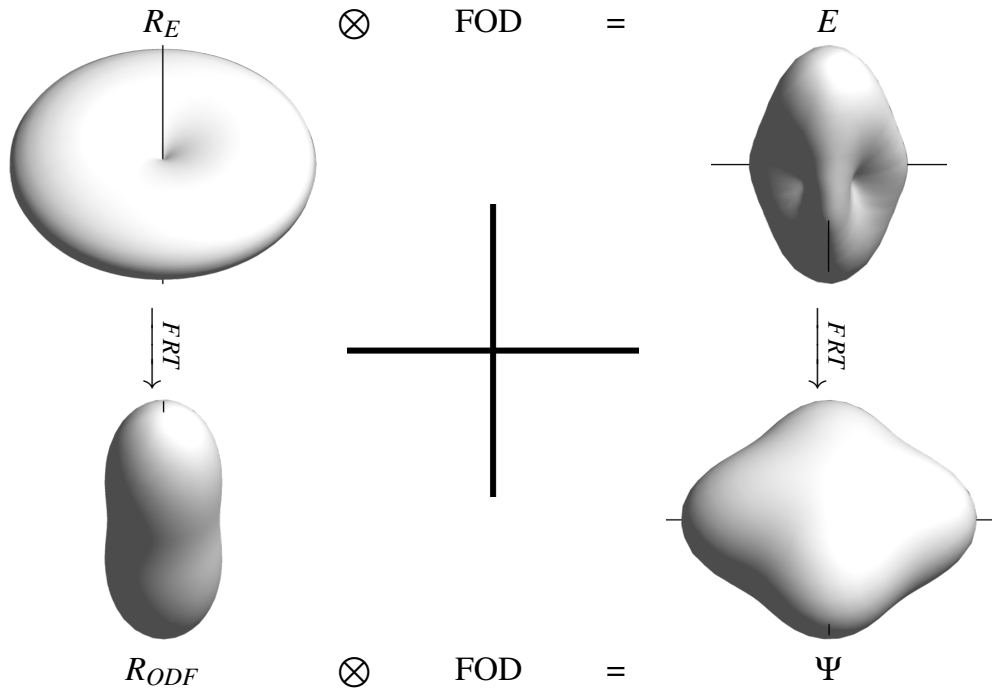


Figure 2: Convolution of a single fibre response with FOD (top – HARDI, bottom – ODF), 8th order SHS and $b = 3000$ mm^2/s.

with $P_l(0)$ being the associated Legendre function of order l evaluated at 0.

The ODF should be relatively smooth with a few maxima oriented along the direction of underlying fibres. Unfortunately, due to noise the ODF has a lot of sharp spikes and needs to be smoothed. Noise related peaks can be reduced by filtering SH coefficients [27], including a regularization scheme in the signal approximation [10, 16], or by selectively removing the noise-infested basis functions [20]. In the first two cases, the smooth ODF function is produced at the cost of a lower angular resolution.

3 Fibre orientation density

Fibre orientation density (FOD) and fibre orientation distribution function (FODF) is a sharper version of ODF. The spherical deconvolution introduced by Tournier [27] allows to compute FOD directly from HARDI data. The measured signal E is a convolution of unknown FOD F with the signal R_E coming from a single fibre population (Figure 2, top):

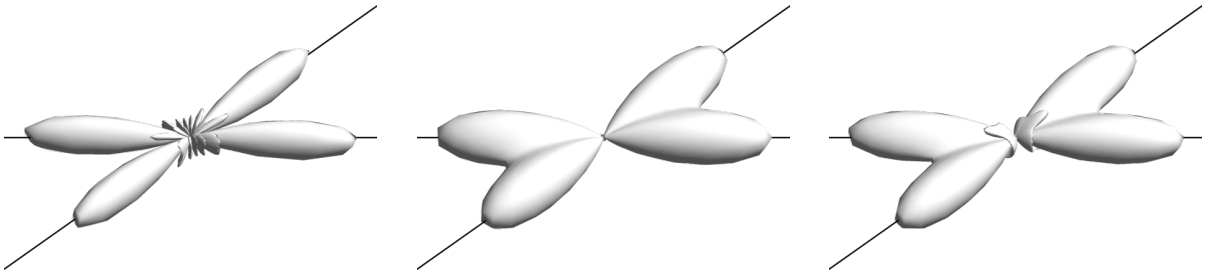
$$E = F \otimes R_E. \tag{6}$$

The single fibre response R_E is either approximated from the most anisotropic voxels [27] or on a voxel-by-voxel basis [2], and represented by rotational harmonics [14]. Since the SHT is a Fourier transform (on the sphere) the convolution can be efficiently represented with a matrix multiplication, or linear transformation of SH coefficients:

$$f_l^m = c_l^m / r_l, \tag{7}$$

where f_l^m is a FOD spherical harmonic coefficient of l order and m band, and r_l a rotational harmonic coefficient of a single fibre response R_E.

Ideally, with an infinite SH series the signal would deconvolve to a sum of delta functions oriented along the underlying fibre tracts (Figure 2, FOD). But as the number of samples is



a) b) c)

Figure 3: Spherical deconvolution to a delta function (a), cosine power lobe (b) and spherical deconvolution transform (c), simulated 45° fibre crossing, 8th order SHS, $b = 3000$ mm^2/s.

limited, the series expansion of the signal has to be truncated which results in a smoother FOD and introduces unwanted ringing near the centre of the lobe.

To remove the ringing, and partially reduce the false FOD peaks that are caused by noise Schultz proposed deconvolution using a non-ringing cosine power lobe ($\cos^h \gamma$, where h depends on the size of SHS used) [25]. The resulting FOD is less sharp (Figure 3b) than the classical deconvolution to a delta function (Figure 3a) [27].

Another way of acquiring FOD is based on sharpening the diffusion ODF. Using the same spherical deconvolution method, it is possible to deconvolve a diffusion ODF to FOD using a single fibre response ODF (Figure 2, bottom). Descoteaux provided a formal relationship between ODF and FOD (called fibre ODF, or FODF, as it was derived from ODF) [11]. The method, called spherical deconvolution transform (SDT), like cosine power lobe, produces less sharp but more noise resilient FOD but, unlike cosine power lobe, introduces a regular negative ringing (Figure 3c).

A different approach was sought by Kezele [18], who reconstructed the sharp diffusion ODF by incorporating a spherical wavelet transform into the Funk–Radon transform. Also Tristan-Vega [29, 30] modified the Funk–Radon approximation to the radial integral. By including the Jacobian of the spherical coordinates in FRT he computed a true orientation probability density function (OPDF). Similar approach, but with a different orientation function (both with and without SHT) was proposed by Özarslan [23] in a diffusion orientation transform (DOT).

FOD has the same maxima as ODF, but the function itself should be sharp. The same methods mentioned in the ODF regularization/smoothing can be used to reduce noise related false peaks, but cannot guarantee the non-negativity of the FOD function. The improved deconvolution algorithms proposed by Dell'Acqua [7], Sakaie [24], and Tournier [28] based on iterative approach (with Dell'Acqua using a modified Richardson-Lucy deconvolution [6], and Sakaie and Tournier a spherical deconvolution) address this.

Finally, it is important to note that the spherical harmonic basis functions are globally supported and thus are not well suited to describe sharp FODs. A recent study by Michailovich, in which the HARDI signal (and subsequently ODF) is modelled using multiresolution bases of spherical ridgelets [19] can match the SH-based QBI (45 basis functions) accuracy-wise with just a few basis functions (4 to 8).

4 Examples

Figure 4 shows diffusion signal, ODF, FODF, and two FODs (delta and cosine power lobe) computed over the same region of interest of a whole-brain scan of a healthy male subject.

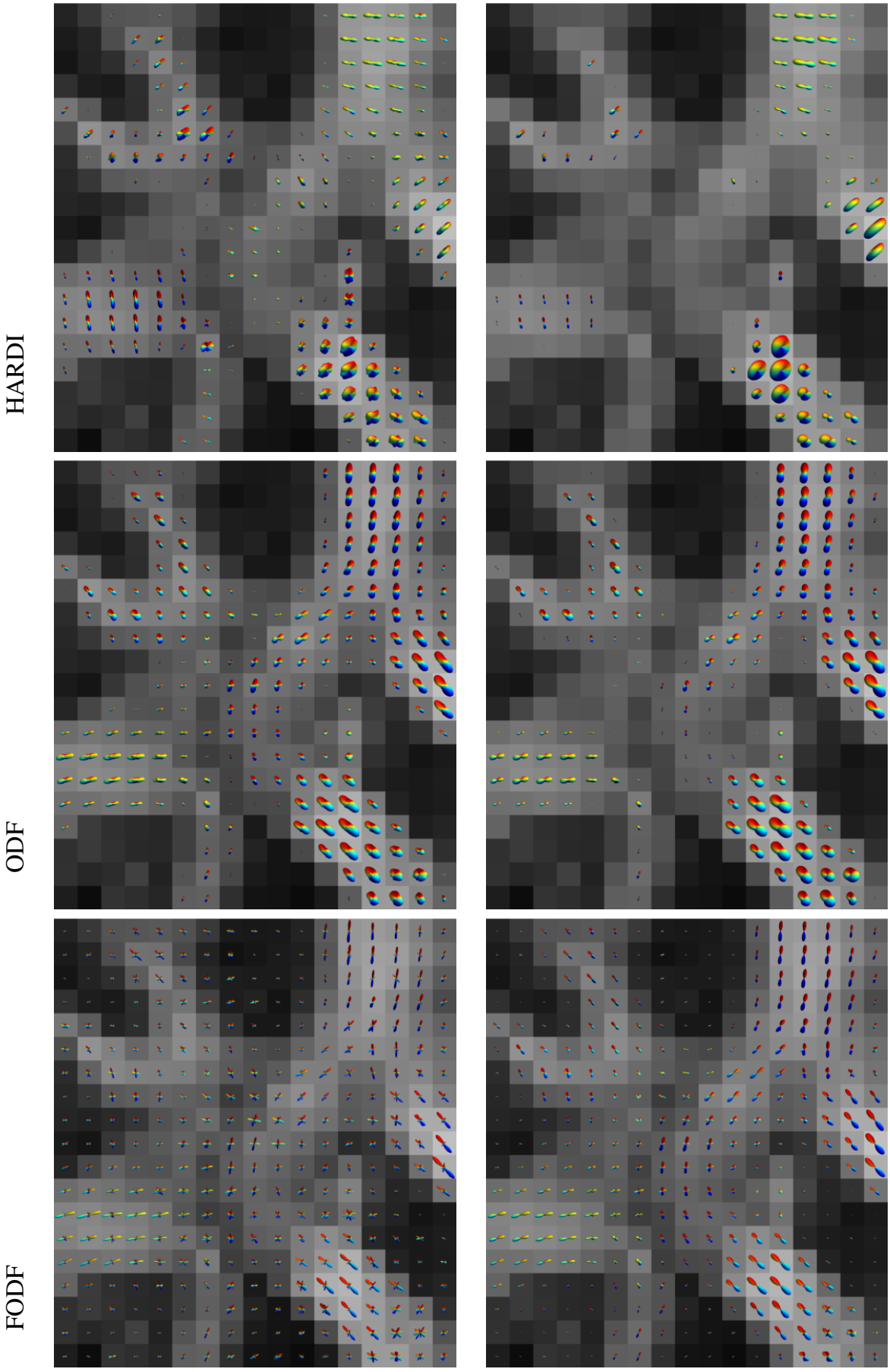


Figure 4

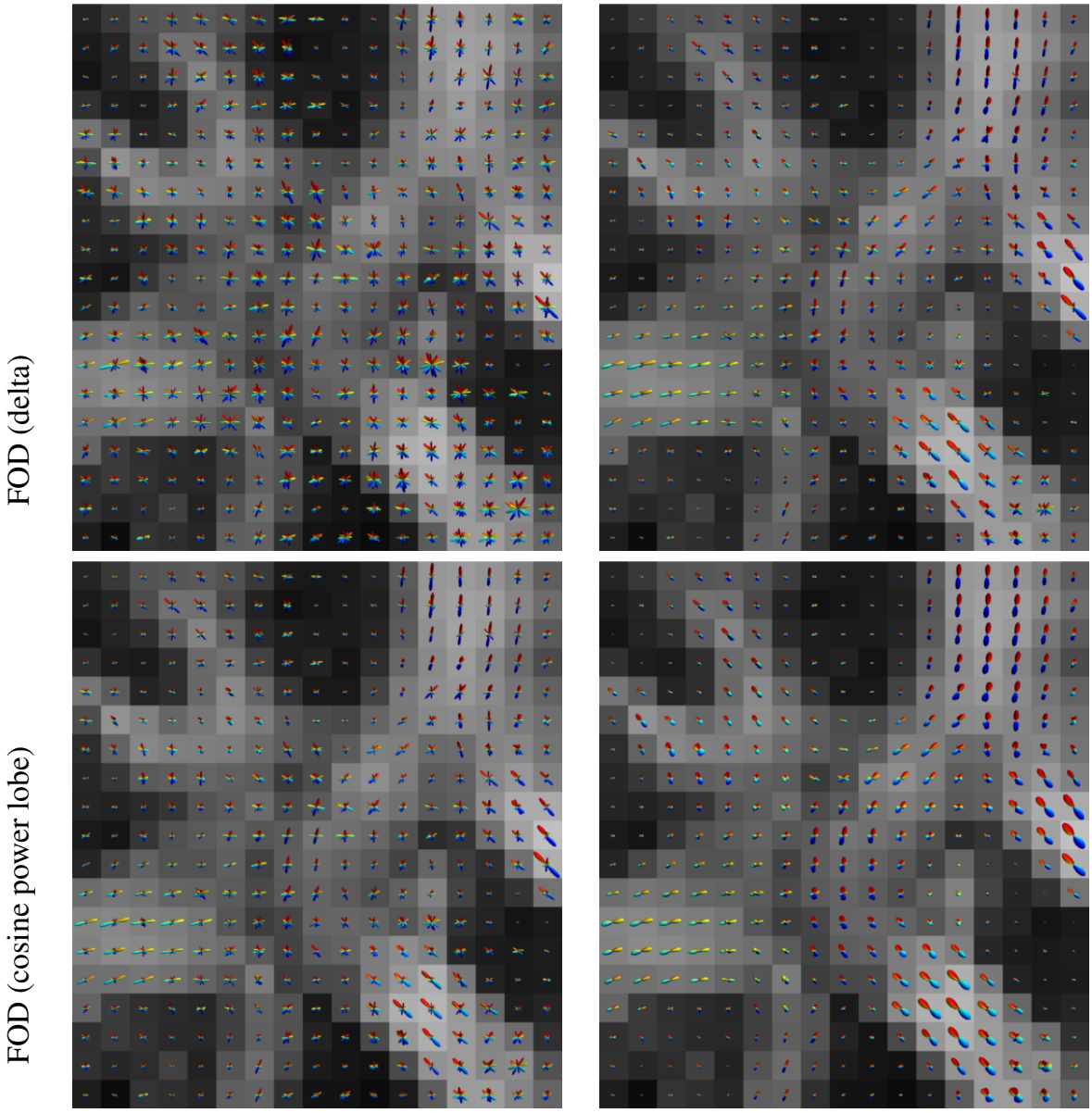


Figure 4: HARDI signal, ODF, FODF, and delta and cosine power lobe FOD profiles plotted over FA of a small region of interest. First column – least squares expansion, second column – regularized expansion using Laplace–Beltrami smoothing operator ($\lambda = 0.05$).

The image was obtained using a single-shot, spin-echo, echo-planar, diffusion-weighted sequence in a Philips 3T Achieva clinical imaging system[1]. Both least squares and smoothed (using Tikhonov regularization with Laplace–Beltrami operator [8]) profiles are provided.

5 Conclusion

Every year new papers related to the processing of diffusion weighted images are published. Methods for computing ODF and FOD has been studied for over 10 years now. This review summarises recent development and provides some insight into the current state of the art in

[1]Acquisition matrix 112x112 with in-plane resolution 2×2 mm^2; 52 slices with a thickness of 2 mm; $b = 3000$ s/mm^2; $TE = 72$ ms; $TR = 15292$ ms; 61 evenly spaced diffusion weighting directions [5]; six $b = 0$ s/mm^2 images acquired and averaged

that broad field. Despite all the work summarized here, the ODF and FOD computation has not yet reached its maturity and still offers a great opportunity for research.

References

[1] D.C. Alexander, G.J. Barker, and S.R. Arridge. Detection and modeling of non-Gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48:331–340, 2002.

[2] A.W. Anderson. Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 54:1194–1206, 2005.

[3] P.J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi. In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, 44:625–632, 2000.

[4] D.L. Bihan. Molecular diffusion, tissue microdynamics and microstructure. *NMR in Biomedicine*, 8:375–386, 1995.

[5] P.A. Cook, M. Symms, P.A. Boulby, and Alexander D.C. Optimal acquisition orders of diffusion-weighted MRI measurements. *Journal of Magnetic Resonance*, 25:1051–1058, 2007.

[6] M.E. Daube-Witherspoon and G. Muehllehner. An iterative image space reconstruction algorthm suitable for volume ECT. *IEEE Transactions on Medical Imaging*, 5:61–66, 1986.

[7] F. Dell'Acqua, G. Rizzo, P. Scifo, R.A. Clarke, G. Scotti, and F. Fazio. A model-based deconvolution approach to solve fiber crossing in diffusion-weighted MR imaging. *IEEE Transactions on Biomedical Engineering*, 54:462–472, 2007.

[8] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Apparent diffusion coefficients from high angular resolution diffusion imaging: estimation and applications. *Magnetic Resonance in Medicine*, 56:395–410, 2006.

[9] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. A fast and robust ODF estimation algorithm in q-ball imaging. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2006.

[10] M. Descoteaux, E. Angelino, S. Fitzgibbons, and R. Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, 58:497–510, 2007.

[11] M. Descoteaux, R. Deriche, T.R. Knosche, and A. Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transactions on Medical Imaging*, 28:269–286, 2009.

[12] G. E. Fasshauer and L. L. Schumaker. Scattered data fitting on the sphere. In *International Conference on Mathematical Methods for Furves and Surfaces II*, 1998.

[13] L.R. Frank. Anisotropy in high angular resolution diffusion-weighted MRI. *Magnetic Resonance in Medicine*, 45:935–939, 2001.

[14] D.M. Healy, H. Hendriks, and P.T. Kim. Spherical deconvolution. *Journal of Multivariate Analysis*, 67:1–22, 1998.

[15] C.P. Hess, P. Mukherjee, E.T. Han, D. Xu, and D.B. Vigneron. A spherical harmonic approach to q-ball imaging. In *International Society for Magnetic Resonance in Medicine*, 2005.

[16] C.P. Hess, P. Mukherjee, E.T. Han, D. Xu, and D.B. Vigneron. Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis. *Magnetic Resonance in Medicine*, 56:104–117, 2006.

[17] K.M. Jansons and D.C. Alexander. Persistent angular structure: new insights from diffusion magnetic resonance imaging data. *Inverse Problems*, 19:1031, 2003.

[18] I. Kezele, M. Descoteaux, C. Poupon, F. Poupon, and J.F. Mangin. Spherical wavelet transform for ODF sharpening. *Medical Image Analysis*, 14:332–342, 2010.

[19] O. Michailovich and Y. Rathi. On approximation of orientation distributions by means of spherical ridgelets. *IEEE Transactions on Image Processing*, 19:461–477, 2010.

[20] B.P. Neuman, L. Bai, and C. Tench. Reliably estimating the diffusion orientation distribution function from high angular resolution diffusion imaging data. In *Medical Image Understanding and Analysis*, 2011.

[21] A.F. Nikiforov, V.B. Uvarov, and R.P. Boas. *Special functions of mathematical physics*. Birkhäuser BaselBoston, 1988.

[22] E. Özarslan and T.H. Mareci. Generalized diffusion tensor imaging and analytical relationships between diffusion tensor imaging and high angular resolution diffusion imaging. *Magnetic Resonance in Medicine*, 50:955–965, 2003.

[23] E. Özarslan, T.M. Shepherd, B.C. Vemuri, S.J. Blackband, and T.H. Mareci. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage*, 31:1086–1103, 2006.

[24] K.E. Sakaie and M.J. Lowe. An objective method for regularization of fiber orientation distributions derived from diffusion-weighted MRI. *NeuroImage*, 34:169–176, 2007.

[25] T. Schultz and H.-P. Seidel. Estimating crossing fibers: A tensor decomposition approach. *IEEE Transactions on Visualization and Computer Graphics*, 14:1635–1642, 2008.

[26] S.N. Sotiropoulos, L. Bai, and C.R. Tench. Fuzzy anatomical connectedness of the brain using single and multiple fibre orientations estimated from diffusion MRI. *Computerized Medical Imaging and Graphics*, 34:504–513, 2010.

[27] J.D. Tournier, F. Calamante, D.G. Gadian, and A. Connelly. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, 23:1176–1185, 2004.

[28] J.D. Tournier, F. Calamante, and A. Connelly. Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, 35:1459–1472, 2007.

[29] A. Tristán-Vega, C.F. Westin, and S. Aja-Fernández. Estimation of fiber orientation probability density functions in high angular resolution diffusion imaging. *NeuroImage*, 47:638–650, 2009.

[30] A. Tristán-Vega, C.F. Westin, et al. A new methodology for the estimation of fiber populations in the white matter of the brain with the Funk-Radon transform. *NeuroImage*, 49:1301–1315, 2010.

[31] D.S. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52:1358–1372, 2004.

[32] D.S. Tuch, T.G. Reese, M.R. Wiegell, N. Makris, J.W. Belliveau, and V.J. Wedeen. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, 48:577–582, 2002.

[33] D.S. Tuch, T.G. Reese, M.R. Wiegell, and V.J. Wedeen. Diffusion MRI of complex neural architecture. *Neuron*, 40:885–895, 2003.

Poster Session 2

Accuracy of prospective motion correction in MRI using tracking markers on repositionable dental impressions

Daniel Stucht[1]
daniel.stucht@ovgu.de
Peter Schulze[1]
peter.schulze@ovgu.de
K. A. Danishad[1]
danishad.appu@ovgu.de
Ilja Y. Kadashevich[1]
ilja.kadashevich@ovgu.de
Maxim Zaitsev[2]
maxim.zaitsev@uniklinik-freiburg.de
Brian S. R. Armstrong[3]
bsra@uwm.edu
Oliver Speck[1]
oliver.speck@ovgu.de

[1] Biomedical Magnetic Resonance
Otto-von-Guericke University
Magdeburg, Germany

[2] Department of Radiology
University Medical Center Freiburg
Freiburg, Germany

[3] Department of Electrical Engineering
University of Wisconsin-Milwaukee
Milwaukee, WI, USA

Abstract

Magnetic resonance imaging (MRI) is a non invasive tool for clinical diagnosis and neuroscience to examine the anatomy of the human brain. Functional MRI (fMRI) even allows studying the neural activity. MRI at ultra high field MRI, such as 7T, offers the possibility to acquire high-resolution MR images. Unfortunately, higher resolution requires longer measurement times, which makes the scans particularly prone to motion artefacts, as motion is more likely to occur over longer scan periods. By using prospective motion correction, artefacts due to patient motion during the measurement can be avoided. If a marker based tracking system is used, automatic registration of multiple scans taken on different days is possible, if the marker can be attached to the subject at the exact same location for every scan. Markers on dental impressions offer this possibility, because they are individually manufactured to match the subject's teeth, which allows a precise repositioning in the upper jaw. This study examines the accuracy of automatic registration of MRI scans with prospective motion correction using an optical tracking system and a passive marker mounted on a dental impression.

1 Introduction

It is a well known problem in clinical and neuroscientific magnetic resonance imaging (MRI) that patient motion during an MRI measurement causes artefacts like blurring and ringing, which reduce the effective resolution of the data and might render the images useless. To exploit the higher SNR of ultra high field (UHF) systems such as 7T for high resolution imaging, longer scan times are necessary, which makes the appearance of patient motion more likely. Typical scan times for high resolution scans (0.4-0.6 mm) are 10-30 minutes and can easily reach several hours for very high resolution MRI (0.1-0.4 mm) of a full brain volume. Even for trained volunteers, it is impossible to remain motionless for so long.

Prospective Motion correction in MRI not only allows correcting for subject motion to avoid artefacts during the measurement; by activating inter-scan motion correction, different scans of the same subject can be aligned to each other automatically. This is useful for follow-up examinations or long time studies with multiple scans, which are usually realigned retrospectively. Inter-scan prospective motion correction also offers new applications such as registering scans from different imaging systems (e.g. MR, CT, PET) if they are equipped with a motion correction system. It would also allow exact repositioning for radiation treatment planning or surgery planning.

If necessary, a scan could also be paused, and could be continued later. Currently, interrupted scans have to be repeated. With this option, very long scans can be performed by acquiring the data in several short rather than in one very long session.

When optical tracking is used, this technique requires the ability to reposition a tracking marker to the same location on the subject relative to the scan volume for every scan. As the upper jaw is rigidly connected to the skull, a dental impression tightly fixed to the subject's teeth offers this possibility. Several systems for prospective motion correction using optical tracking systems such as infrared based stereoscopic tracking of retro reflective markers [3, 7], single camera based systems based on moiré phase tracking [2, 5] or tracking using features like 2D patterns [1] have been presented. In these studies the markers are attached to the subject on goggles [1, 2] or are attached to the subject's skin at the forehead [5] Both options make exact repositioning very difficult if not impossible. Dold *et al.* [3] and Zaitsev *et al.* [7] have used dental impressions, but as they didn't do inter-scan alignment, exact repositioning was not required.

This study examines the accuracy of intra-subject registration by prospective motion correction using an optical tracking system and a tracking marker on a dental impression.

2 Materials and methods

This section describes the data acquisition, the motion correction system and the analysis of the image data. The calculation of the residual registration error was performed using a Matlab (The MathWorks, Natick, MA, USA) implementation of statistical parametric mapping (SPM8, Wellcome Trust Centre for Neuroimaging, UCL, London, UK).

2.1 Image acquisition

MRI Measurements were performed on a 7T whole body MRI (Siemens Medical Solutions, Erlangen, Germany) using a 32-channel coil (Nova Medical, Wilmington, MA, USA) and the following sequence parameters for two subjects: imaging matrix = 224 × 224 × 72,

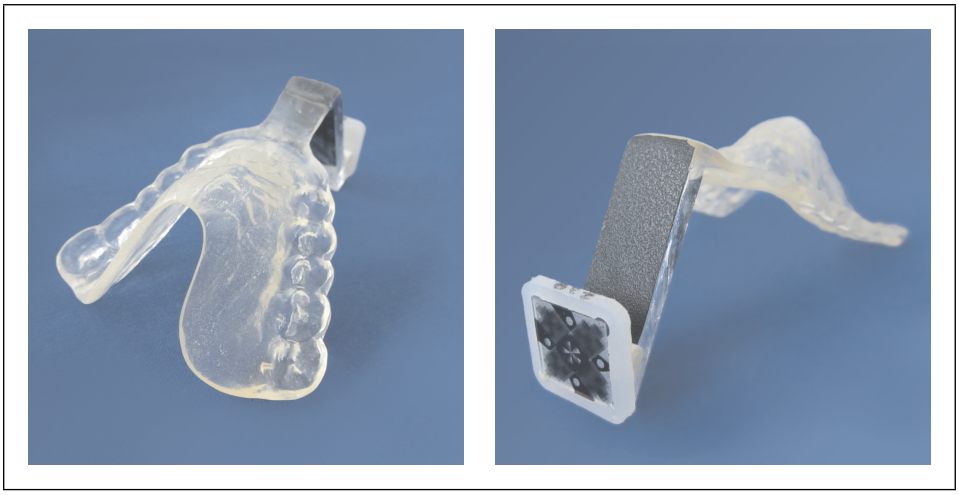


Figure 1: The dental impression (left) individually created for one of the subjects and the MPT marker attached to the retainer (right).

voxel size = 1.0 × 1.0 × 1.0 mm^3, flip angle α = 5°. The repetition time and echo time (TR/TE) were different for both subjects: subject 1: TR/TE = 9.8 ms/3.38 ms; subject 2: TR/TE = 14.0 ms/9.0 ms. For each subject the scan was repeated 5 times. Between every scan, the subjects were removed from the scanners bore, the head coil was opened, the subject moved the head out of coil to remove the mouthpiece and repositioned it. The position of the head in the coil was not observed, there were no specific actions taken for an exact repositioning of the head. This procedure and the preparation for the next scan took approx. two minutes, which was chosen as the time between two scans.

2.2 Prospective motion correction and volume alignment

Schulze *et al.* [6] have tested three different optical tracking systems. A single-camera in-bore system based on moiré phase tracking (MPT) (University of Wisconsin-Milwaukee, Milwaukee, WI, USA) showed best accuracy and was used to accomplish the motion tracking in the study presented here. The standard deviation of position-data of the X, Y- and Z- axes are below 4 µm.

The tracking marker was attached to a small retainer with a dental impression (Figure 1), which was individually manufactured for both subjects and tightly fixed to the subject's upper jaw.

The coordinate systems of the tracking system and the MRI scanner were carefully co-registered prior to the MR session by using a non-iterative cross-calibration algorithm described by Kadashevich *et al.* [4].

At the beginning of the first scan, the position of the marker (i.e. the dental impression) relative to the camera of the tracking system was saved as reference. For the four follwing scans, the scan volume was adjusted to match this reference position by recalculating the scanners gradients and frequencies in real time. Thus the differences in marker positions between the first and the subsequent scans were corrected during the measurement, which enables intra- and inter-scan motion correction.

2.3 Calculation of residual registration error

To calculate the residual registration error, the data were processed using the realign function of SPM8. The first scan was taken as a reference, scans two to five were aligned to the first scan. Pre-processing the data by performing a brain extraction with the Brain Extraction Tool (BET) of FSL 4.1.2 (www.fmrib.ox.ac.uk/fsl) did not show any significant change in the error calculation, thus the data shown here were calculated on the full MRI volumes.

3 Results

One slice from each of the five volumes taken from subject 1 is shown in Figure 2. Differences are hardly noticeable. Figure 3. shows the result of the SPM realign calculations with a residual error clearly below one millimetre for translation on all three axes. The total length of the translation vector is shown in table 1. The rotational components differ in the range of approx. 1 degree. It is noteworthy that these values represent the error of the whole motion correction system. This includes the noise of the tracking data, residual errors in the cross calibration between tracking system and scanner and the misplacements of the mouthpiece.

4 Discussion

The results show that automatic alignment of intra-subject scans is possible with a good accuracy. The misalignment is below the level of motion often observed at patient scans. With this method, clinical scans that had to be interrupted for any reason, could be paused and continued with only minor artefacts. This would reduce the need for repeated scans and thus improve clinical workflow, make the examination more convenient and save costs.

For interrupted high resolution imaging, where a single scan does not generate a full data set due to the long scan times, additional retrospective registration might be necessary to correct for the residual misalignment. This will be much easier, as the technique ensures that the volume of interest will be almost completely covered by multiple successive scans. The evaluation of these applications will be subject to following studies.

A better separation of the different error sources (tracking noise, imperfect cross calibration, misplaced marker) would be desirable. This could be accomplished by comparing the results of this study to those from scans consecutively taken with motion between the scans, but without repositioning of the dental impression.

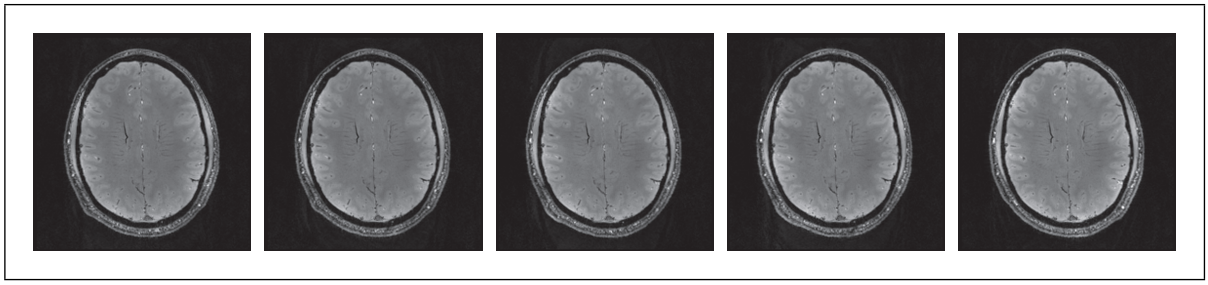


Figure 2: The same slice from the five scans taken from subject 1 in chronological order (left to right). The image on the left shows the reference scan.

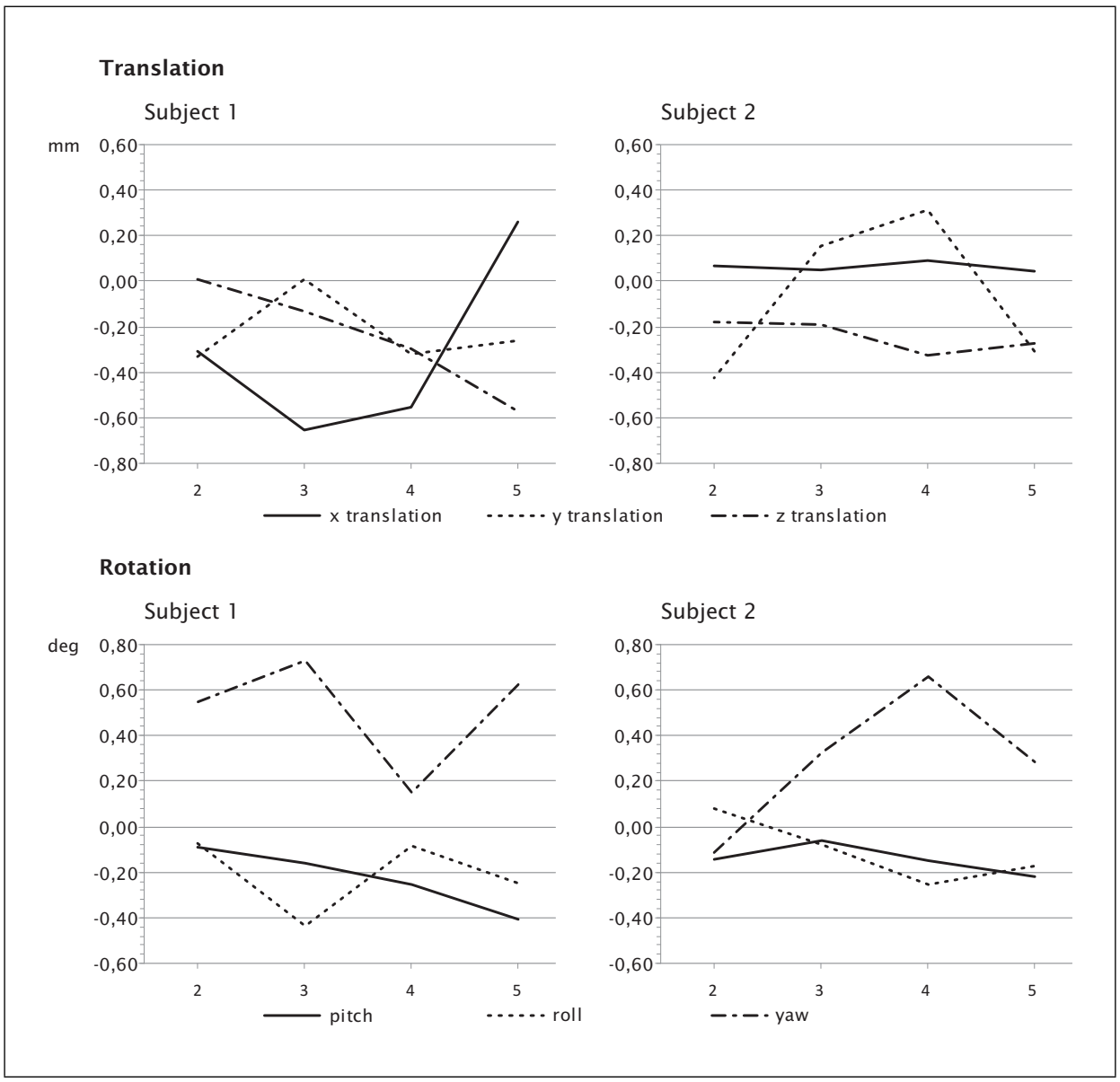


Figure 3: Residual registration error calculated by SPM8. The graphs show the calculated difference between scans 2 to 4 and the first scan. Errors in x-, y- and z-translation (upper row) and pitch, roll and yaw rotations (lower row) for two subjects.

		Subject 1	Subject 2
Total 3D translation (mm)	Mean	0.63	0.40
	Standard deviation	0.12	0.10
pitch (deg)	Mean	-0.23	-0.14
	Standard deviation	0.14	0.07
roll (deg)	Mean	-0.21	-0.11
	Standard deviation	0.17	0.14
yaw (deg)	Mean	0.51	0.29
	Standard deviation	0.25	0.32

Table 1: Mean and standard deviation of the translational (3D vector length) and rotational components of the registration error.

References

[1] Murat Aksoy, Christoph Forman, Matus Straka, Stefan Skare, Samantha Holdsworth, Joachim Hornegger, and Roland Bammer. Real-time optical motion correction for diffusion tensor imaging. *Magn Reson Med*, 66(2):366–378, 2011.

[2] Brian C. Andrews-Shigaki, Brian S.R Armstrong, Maxim Zaitsev, and Thomas Ernst. Prospective motion correction for magnetic resonance spectroscopy using single camera retro-grate reflector optical tracking. *J Magn Reson Imaging*, 33(2):498–504, 2011.

[3] Christian Dold, Maxim Zaitsev, Oliver Speck, Evelyn Firle, Jürgen Hennig, and Georgios Sakas. Prospective head motion compensation for mri by updating the gradients and radio frequency during data acquisition. In James Duncan and Guido Gerig, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005*, volume 3749 of *Lecture Notes in Computer Science*, pages 482–489. Springer, Berlin / Heidelberg, 2005.

[4] Ilja Kadashevich, Appu Danishad, and Oliver Speck. Automatic motion selection in one step cross-calibration for prospective mr motion correction. *Proc. Magn Reson Mater Phy*, 24(Supplement 1):266–267/#371, 2011.

[5] Julian Maclaren, Brian S. Armstrong, Robb T. Barrows, A. K. Danishad, Thomas Ernst, Colin L. Foster, Kazim Gumus, Michael Herbst, Ilja Y. Kadashevich, Todd P. Kusik, Qiaotian Li, Cris Lovel-Smith, Tom Prieto, Peter Schulze, Oliver Speck, Daniel Stucht, and Maxim Zaitsev. Measurement of microscopic head motion during brain imaging. In *Proceedings of the 20th Scientific Meeting: International Society for Magnetic Resonance in Medicine (ISMRM 2012)*, page #373, 2012.

[6] Peter Schulze, Ilja Kadashevich, Daniel Stucht, Appu Danishad, and Oliver Speck. Prospective motion correction at 7 tesla magnetic resonance imaging using optical tracking systems. In *10. Magdeburger Maschinenbau-Tage "Forschung und Innovation"*, pages C1–3, Magdeburg, 2011.

[7] Maxim Zaitsev, Christian Dold, Georgios Sakas, Jürgen Hennig, and Oliver Speck. Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system. *Neuroimage*, 31(3):1038–1050, 2006.

GPF Deformable Model based Vessel Segmentation in CT

Si Yong Yeo[1]
genyeoius@gmail.com
Xianghua Xie[2]
x.xie@swansea.ac.uk
Igor Sazonov[1]
i.sazonov@swansea.ac.uk
Perumal Nithiarasu[1]
p.nithiarasu@swansea.ac.uk

[1] College of Engineering
Swansea University
Swansea, SA2 8PP, UK

[2] Department of Computer Science
Swansea University
Swansea, SA2 8PP, UK

Abstract

In this work, we present a method to segment carotid vessels from CT scans. Image denoising is performed using vessel enhancing diffusion, which can smooth out image noise and enhance vessel structures. The Canny edge detection technique which produces object edges with single pixel width is used for accurate detection of the lumen boundaries. The image gradients are then used to compute the geometric potential field which gives a global representation of the geometric configuration. The deformable model uses a regional constraint to suppress calcified regions for accurate segmentation of the vessel geometries. The preliminary result shows the prosed method achieves promising results based on qualitative evaluation using manual labelled groundtruth.

1 Introduction

The human circulatory system consists of vessels that transport blood throughout the body, providing the tissues with oxygen and nutrients. It is known that vascular diseases such as stenosis and aneurysms are often associated with changes in blood flow patterns and the distribution of wall shear stress. Modelling and analysis of the hemodynamics in the human vascular system can improve our understanding of vascular disease, and provide valuable insights which can help in the development of efficient treatment methods. One of the main challenges is the accurate reconstruction of the vascular geometry. The anatomical information used to reconstruct the geometric models are usually provided in the form of medical image datasets (scans) from imaging modalities such as computed tomography (CT) and magnetic resonance (MR) imaging. Manual reconstruction of the vasculature geometries can be tedious and time consuming. There is also the issue of variability between the geometries extracted manually by different individuals, and variability of geometries extracted by the same individual at different occasions.

Although several techniques exist for the segmentation of vascular structures from medical images, it remains an intricate process due to factors such as image noise, partial volume

effects, image artifacts, intensity inhomogeneity and changes in topology. In [11], the coordinate points for the center line of the aortic arch were extracted from volume rendered MR images. A cubic spline was then used to represent the aortic centerline, and cross-sectional grids were generated on normal planes at equidistant points along the curve. This generated a curved tube with circular cross section of uniform radius, which is not representative of the geometry of the aorta. In [15], the centerline and diameter information of the vessels was extracted from the image dataset, and the vascular model was reconstructed using non-uniform rational B-splines (NURBS). Such techniques may often smooth out geometric information that can be important to the computation of accurate flow dynamics, such as those at bifurcations.

The 3D models of the vascular structures are commonly reconstructed by extracting the 2D contours of the vessels at each of the image slices, and then lofting through the contours to create the surface models of the vessels, e.g. [3, 16, 19]. In [8, 9], a 3D dynamic surface model was used to delineate the boundary of carotid arteries. An initial triangulated model was placed within the interior of the carotid vessels, and an inflation force was applied to deform the model towards the vessel wall. In particular, the inflation force is applied only when the vertices of the model are within the lumen, i.e., at locations with image intensity below a user-specified threshold. An image-based force is further applied to the surface model to better localize the boundary. It may however be difficult to select an appropriate threshold value that delineates the vessel wall closely due to inhomogeneous image intensity. This approach is sensitive to noise, and manual editing is often required to move the vertices towards the vessel wall. In [13], a 2D discrete dynamic contour was first used to extract the vessel contours, a dynamic surface model was then inflated to reconstruct the surface model using the binary images of the extracted contours. This however does not consider the 3D geometric information from the image dataset. In [4, 5, 18], the surface models for each of the vessel branches of the carotid artery were reconstructed independently using a tubular deformable model. A surface merging algorithm is then required to reconstruct the surface model of the carotid bifurcation from the triangulated surfaces of the vessel branches. This particular approach requires the determination of the axis of each of the vessels, which can be done manually by selecting a reasonable amount of points from image slices to represent the curves of the structure. Due to the smoothing effect of this technique, regions of high curvature such as those at bifurcations or stenosis may not be modeled accurately. These explicit deformable models represent contours and surfaces parametrically, which requires the tracking of points on the curves and surfaces during deformation. It is therefore difficult for explicit deformable models to deal with topological variation and complex shapes.

Implicit deformable models have been applied in the segmentation of vascular structures in [1, 2, 6, 12, 14]. However, many of these techniques use an attraction force field which acts on contours or surfaces only when they are close to the object boundaries. As such, initial contours have to be placed close to the object boundaries, which can be tedious in complex geometries. A constant pressure term such as the one in [10], is often used to monotically expand or shrink the deformable model towards the image object boundaries, which can overwhelm weak object edges. In addition, the initial contours have to be placed either inside or outside object boundaries, which can be difficult for compact and narrow structures. Many of these techniques are also sensitive to image noise, and have difficulties in extracting deep boundary concavities.

2 Proposed Method

We propose a 3D deformable model based segmentation method to extract the carotid structures. The carotid vessels are first enhanced by applying anisotropic diffusion, and then vessel edges are localised using Canny edge detection which provides better edge localisation and connectivity compared to other conventional techniques, such as Sobel. The images gradient vectors at those edge locations identified by Canny edge detector are then used to compute a geometric potential field which gives a global representation of the geometric configuration. This field is then used to drive a 3D deformable model, with an additional simplistic regional constraint to suppress the interference from vessel calcificcation.

2.1 Vessel enhancing and edge detection

The formulation of the vessel enhancing diffusion filter is based on a smoothed version of the vesselness measure used in [7]. In this approach, an anisotropic diffusion filter with strength and direction determined by the vesselness measure is applied to enhance the geometric structures of the vessel. The vesselness measure is determined by analyzing the eigensystem of the Hessian matrix given as:

$$\mathbf{H} = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{bmatrix} \tag{1}$$

which describes the geometric information at each point of a 3D image I based on the local intensity variations. Here, the derivatives of the image I are computed as convolution with derivatives of the Gaussian function, i.e. $I_x = I(\mathbf{x}) * \frac{\partial}{\partial x} G_\sigma(\mathbf{x})$, where G_σ denotes the Gaussian function with standard deviation σ. The principal curvatures and directions are given by the maximum and minimum eigenvalues and the corresponding eigenvectors. With the eigenvalues given such that $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$, the vesselness measure is defined as: if $\lambda_2 \geq 0$ or $\lambda_3 \geq 0$, $V_\sigma(\lambda) = 0$; otherwise $V_\sigma(\lambda) = \left(1 - e^{\frac{-{R_A}^2}{2\alpha^2}}\right) \cdot e^{\frac{-{R_B}^2}{2\beta^2}} \cdot \left(1 - e^{\frac{-S^2}{2\gamma^2}}\right) \cdot e^{\frac{2c^2}{|\lambda_2|{\lambda_3}^2}}$ with $R_A = \frac{|\lambda_2|}{|\lambda_3|} R_B = \frac{|\lambda_1|}{\sqrt{|\lambda_2 \lambda_3|}} S = \sqrt{{\lambda_1}^2 + {\lambda_2}^2 + {\lambda_3}^2}$ in which R_A and R_B can be used to differentiate tubular structures from blob-like and plate-like structures, while S is used to differentiate between foreground vessel structures and background noise. The parameters α, β and γ are weighting factors which control the sensitivity of the vesselness measure, and c is a small constant.

For a multiscale analysis, the vesselness function is computed for a range of scales, and the maximum response is selected using the following equation: $V = \max_{\sigma_{min} \leq \sigma \leq \sigma_{max}} V_\sigma(\lambda)$A diffusion tensor is then defined such that vessel diffusion takes place in the direction of the vessel, while diffusion perpendicular to the vessel direction is inhibited. The diffusion tensor can therefore be used to preserve vessel structures and is given as: $\mathbf{D} = \mathbf{Q}\lambda'\mathbf{Q}^T$ where $\mathbf{Q}$ is a matrix containing the eigenvectors of the Hessian matrix $\mathbf{H}$, and λ' is a diagonal matrix with elements given as: ${\lambda_1}' = 1 + (w-1) \cdot V^{\frac{1}{s}}$, ${\lambda_2}' = {\lambda_3}' = 1 + (\varepsilon - 1) \cdot V^{\frac{1}{s}}$ with w, ε and s as tuning parameters. The anisotropic diffusion is then defined as: $L_t = \nabla \cdot (\mathbf{D} \nabla L)$where $L(0)$ is set as the input image.

The Canny edge detection can produce object edges with single pixel width, and can therefore be used for more accurate edge detection of the vessel structures. The image gra-

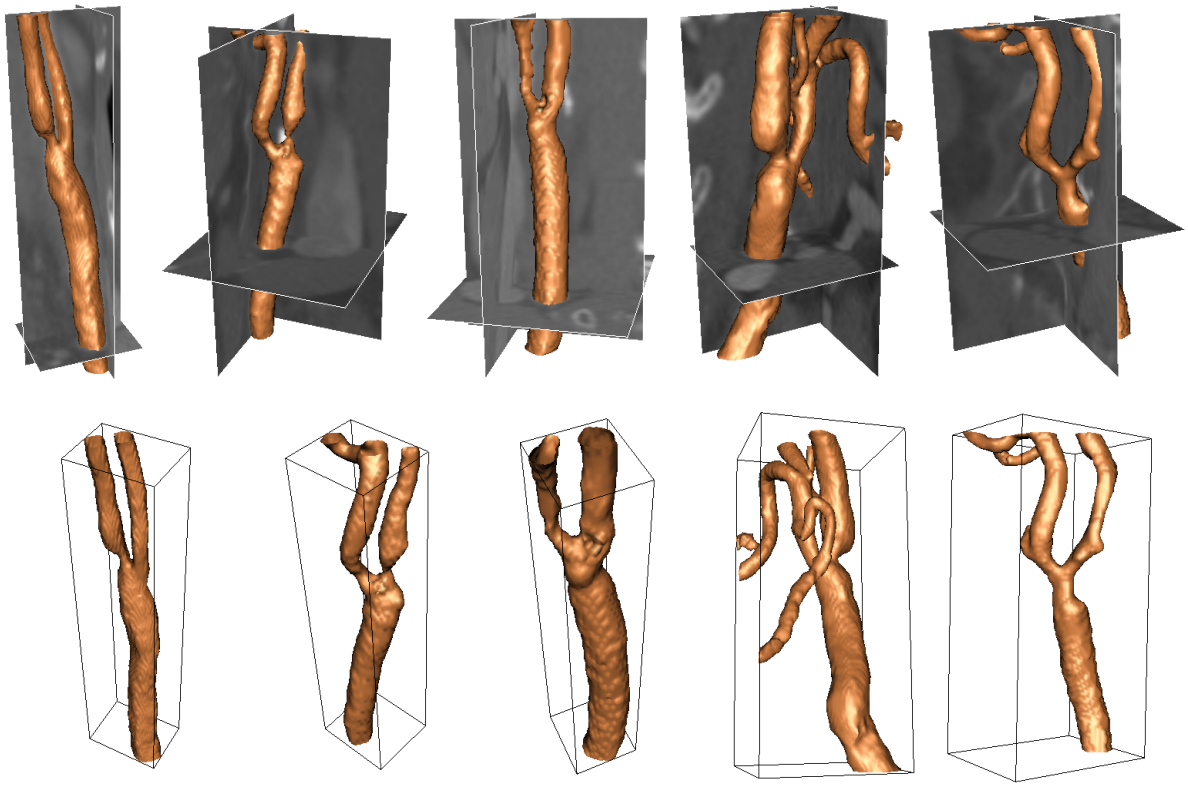


Figure 1: Segmentation of carotid artery from CT image dataset.

dients at the detected edges are then used to compute the geometric potential field, which is briefly described in the following section.

2.2 GPF based segmentation with a simplistic region constraint

It is shown in [17] that the GPF deformable model can be used to efficiently segment complex geometries from biomedical images. By using pixel or voxel interactions across the whole image domain, the deformable model is more robust to image noise and weak edges. The dynamic vector force field changes according to the relative postition and orientation between the geometries, which allows the deformable model to propagate through long tubular structures. Here, the GPF deformable model is applied to segment the geometries of human carotid arteries from CT images. Some of the main challenges in the segmentation of the carotid geometries include intensity inhomgeneity, weak edges and adjacent veins with similar intensities to the carotids. In addition, calcifications which are attached to the arterial walls should not be included in the reconstructed geometries. Although, the calcified plaques often appear as relatively bright regions compared to soft tissues, plaques with lower densities may have similar intensities to the lumen. As the intensities of the plaques vary with the densities, it is not easy for techniques such as global intensity threshold to remove the plaques from the extracted geometries. In this section, a simple heuristic region constraint is added to the deformable model such that it does not propagate across the calcified regions. This is done by constraining the deformable model from propagating across regions with image gradient magnitude larger than a user specified value, T_{max}. As the calcified regions usually have relatively large image gradients, the threshold value can be easily selected by observing the histogram of the image gradient magnitude. The deformable model with

region constraint can thus be expressed as:

$$\frac{\partial \phi}{\partial t} = \begin{cases} 0 & \text{if } |\nabla I| > T_{max} \\ \alpha g \kappa |\nabla \phi| - (1-\alpha)(\mathbf{F} \cdot \nabla \phi) & \text{otherwise} \end{cases} \quad (2)$$

where α is a weighting parameter, g is the edge stopping function, κ is the curvature and $\mathbf{F}$ is the geometric potential force defined in the GPF model [17].

3 Results

In this section, experimental results on the segmentation of the cartoid geometries using the proposed framework are shown. Six datasets from CT imaging are used in the experiment. The volumes of interest containing the carotid arteries are extracted from the image datasets to reduce the size of the input datasets.

Figure 1 shows the segmentation of the carotid geometries using the GPF deformable model with region constraint. The bidirectional and dynamic vector force allows the flexible cross-boundary intializations of the model to easily propagate and converge to the geometries of the carotid arteries. Note that the deformable model easily propagate through the stenotic carotid bifurcations and get around the calcified regions to efficiently segment the carotid geometries from the CT images.

The reconstructed vessel geometries using the proposed method are compared against geometries from manual segmentation. Figure 2 demonstrates the comparison of the extracted geometries using random cross-section slices taken along the z-axis direction. The blue and orange contours represent the cross-section of the geometries extracted manually and using the GPF deformable model respectively. As shown in the figures, the image dataset consist of other tissue structures which may affect the geometric reconstruction. In particular, vessels adjacent to the carotid artery can often cause other models to leak out due to the similar intensity. The geometric potential field provides a more coherent and global representation of the object edges, and allows the deformable model to extract the geometry accurately. By adding a region constraint, the proposed model can reasonably cope with the calcified regions as the deformable model propagates through the tubular structures to segment the vessel geometry. Quantitative analysis based on manual labelling showed promising result. The average foreground accuracy, background accuracy and overall accuracy are 93.9%, 99.8% and 96.8%, respectively. Note, these are normalised accuracy measurement to reduce measurement bias towards the large number of background voxels in the image.

References

[1] L. Antiga and A. Ene-Iordache, B. Remuzzi. Computational geometry for patient-specific reconstruction and meshing of blood vessels from mr and ct angiography. *IEEE T-MI*, 22(5):674–684, 2003.

[2] L. Antiga, M. Piccinelli, L. Botti, B. Ene-Iordache, A. Remuzzi, and Steinman D. A. An image-based modeling framework for patient-specific computational hemodynamics. *Medical and Biological Engineering and Computing*, 46(11):1097–1112, 2008.

[3] A. D. Augst, D. C. Barratt, A. D. Hughes, S. A. McG Thom, and X. Y. Xy. Various issues relating to computational fluid dynamics simulations of carotid bifurcation flow based on models reconstructed from three-dimensional ultrasound images. *Proc Inst Mech Eng H, Journal of Engineering in Medicine*, 217(5):393–403, 2003.

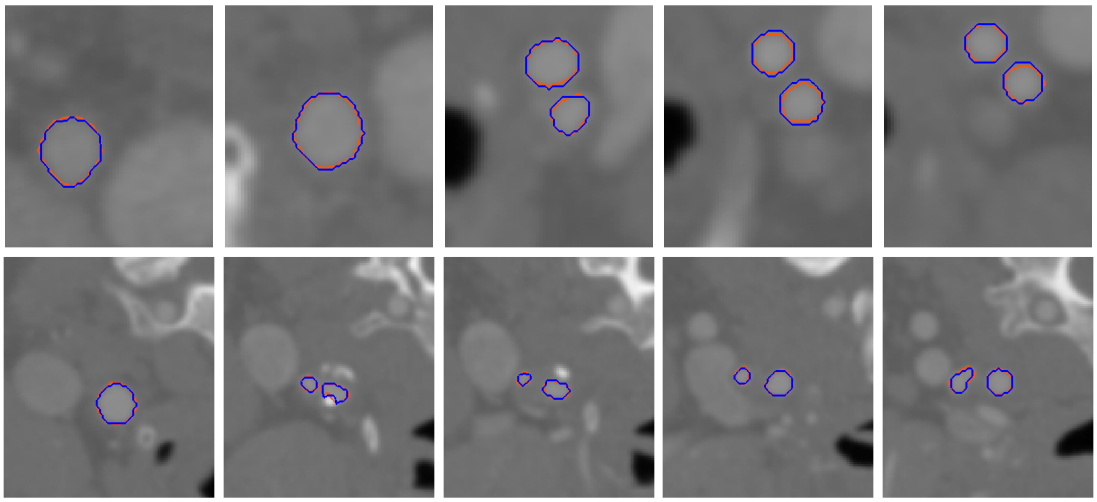


Figure 2: Comparison of geometry segmented from CT image dataset 1 using image slices taken along z-axis direction: blue - manual, orange - GPF deformable model.

[4] J. R. Cebral, R. Lohner, O. Soto, P. L. Choyke, and P. J. Yim. Patient-specific simulation of carotid artery stenting using computational fluid dynamics. In *MICCAI*, pages 153–160, 2001.

[5] J. R. Cebral, M. A. Castro, R. Lohner, J. E. Burgess, R. Pergolizzi, and C. M. Putman. Recent developments in patient-specific image-based modeling of hemodynamics. In *ENIEF04*, 2004.

[6] T. Deschamps, P. Schwartz, D. Trebotich, P. Colella, D. Saloner, and R. Malladi. Vessel segmentation and blood flow simulation using level-sets and embedded boundary methods. In *Computer Assisted Radiology and Surgery*, pages 75–80, 2004.

[7] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In *MICCAI*, pages 130–137, 1998.

[8] J. D. Gil, H. M. Ladak, D. A. Steinman, and A. Frenster. Accuracy and variability assessment of a semiautomatic technique for segmentation of the carotid arteries from three-dimensional ultrasound images. *Medical Physics*, 27(6):1333–1342, 2000.

[9] H. M. Ladak, J. S. Milner, and D. A. Steinman. Rapid three-dimensional segmentation of the carotid bifurcation from serial MR images. *Journal of Biomechanical Engineering*, 122(1):96–99, 2000.

[10] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modelling with front propagation: A level set approach. *IEEE T-PAMI*, 17(2):158–175, 1995.

[11] D. Mori and T. Yamaguchi. Construction of the CFD model of the aortic arch based on mr images and simulation of the blood flow. In *International Workshop on Medical Imaging and Augmented Reality*, pages 111–116, 2001.

[12] B. Nilsson and A. Heyden. A fast algorithm for level set-like active contours. *Pattern Recognition Letters*, 24(9):1311–1337, 2003.

[13] D. A. Steinman, J. B. Thomas, H. M. Ladak, J. S. Milner, B. K. Rutt, and J. D. Spence. Reconstruction of carotid bifurcation hemodynamics and wall thickness using computational fluid dynamics and mri. *Magnetic Resonance in Medicine*, 47(1):149–159, 2002.

[14] J. Svensson, R. Gardhagen, E. Heiberg, T. Ebbers, D. Loyd, T. Lanne, and M. Karlsson. Feasibility of patient specific aortic blood flow CFD simulation. In *MICCAI*, pages 257–263, 2006.

[15] Y. Tokuda, M.-H. Song, Y. Ueda, A. Usui, A. Toshiaki, S. Yoneyama, and S. Maruyama. Three-dimensional numerical simulation of blood flow in the aortic arch during cardiopulmonary bypass. *European Journal of Cardio-thoracic Surgery*, 33(2):164–167, 2008.

[16] X. Y. Xu, Q. Long, M. W. Collins, M. Bourne, and T. M. Griffith. Reconstruction of blood flow patterns in human arteries. *Proc Inst Mech Eng H, Journal of Engineering in Medicine*, 213(5): 411–421, 1999.

[17] S. Y. Yeo, X. Xie, I. Sazonov, and P. Nithiarasu. Geometrically induced force interaction for three-dimensional deformable models. *IEEE T-IP*, 20(5):1373–1387, 2011.

[18] P. J. Yim, J. J. Cebral, R. Mullick, H. B. Marcos, and R. L. Choyke. Vessel surface reconstruction with a tubular deformable model. *IEEE T-MI*, 20(12):1411–1421, 2001.

[19] H. F. Younis, M. R. Kaazempur-Mofrad, R. C. Chan, A. G. Isasi, D. P. Hinton, A. H. Chau, L. A. Kim, and R. D. Kamm. Hemodynamics and wall mechanics in human carotid bifurcation and its consequences for atherogenesis: investigation of inter-individual variation. *Biomechanics and Modeling in Mechanobiology*, 3(1):17–32, 2004.

ICP Variants Robustness to Gaussian and Impulsive Noise

Stuart Henry[1]
henry-r3@email.ulster.ac.uk
Philip Morrow[1]
pj.morrow@ulster.ac.uk
John Winder[2]
rj.winder@ulster.ac.uk
Bryan Scotney[1]
bw.scotney@ulster.ac.uk

[1] School of Computing and Information Engineering, University of Ulster, UK

[2] School of Health Sciences, University of Ulster, UK

Abstract

Iterative closest point (ICP) is a 3D surface-based rigid registration algorithm that locates and uses corresponding point pairs to compute the registration transformation. In the registration of 3D surfaces ICP assumes outlier free data, using all the corresponding point pairs regardless of their distance from the data surface to the closest point in the model surface. Computed tomography (CT) data was acquired from four anatomical phantoms. Two outlier robust variants of ICP, Haralick ICP and Trimmed ICP, were investigated to evaluate their robustness to Gaussian and impulsive noise. The results show that Trimmed ICP produces an accurate registration for both noise types but is computationally expensive.

1 Introduction

Diagnosis and treatment planning of specific diseases requires image acquisition from a range of 3D volumetric modalities including CT and magnetic resonance (MR) [1]. Multimodal data provides structural and functional information adding extra value to each of the modalities involved. Combining image data requires image registration to compute data alignment. ICP is a surface-based rigid registration algorithm that assumes outlier free data, reducing its robustness with real world data [2]. Haralick ICP (HICP) is an outlier robust method that assigns a weight to every corresponding point pair for pose estimation [3]. Corresponding point pairs with the largest residual error are assigned a minimal weighting value to limit impact on pose estimation. Trimmed ICP (TrICP) sorts the distances between each point in the data surface and the closest point in the model surface, selecting a subset of the closest point pairs to compute the registration [4].

In this investigation two CT surfaces, from four anatomical phantoms are registered, with the data surface having different levels of Gaussian and impulsive noise added to simulate outliers. The emphasis of this investigation is to establish the effect that outliers, simulated by Gaussian and impulsive noise, has on the registration of ICP variants.

2 Overview of ICP Variants

2.1 ICP

ICP is a representation independent algorithm that uses point sets to compute rigid registration transformations between two surfaces. The largest volume or highest resolution surface is assigned as the "model" surface (static during registration) with the remaining surface assigned as the "data".

1. Assume that both surfaces are in point sets with the data surface, $\boldsymbol{P}$, having N_p points, $\{p_i,\ i = 1,\ldots,N_p\}$ and the model surface, $\boldsymbol{M}$, having N_m points, $\{m_j,\ i = 1,\ldots,N_m\}$. For each data point p_i find the closest point in **M**:
$$d(p_i, M) = \min_{m \epsilon M} \|m - p_i\| \tag{1}$$
2. Using the correspondence pairs, compute the optimal transformation, rotation ($\boldsymbol{R}$) and translation ($\boldsymbol{T}$) to minimise the point pairs' mean square error (MSE), where $\hat{m}_i$ is the point in $\boldsymbol{M}$ closest to p_i:
$$MSE = \frac{1}{N_p} \sum_{i=1}^{N_p} \|\hat{m}_i - \boldsymbol{R}(p_i) - \boldsymbol{T}\|^2 \tag{2}$$

Optimal transformation estimation uses closed-form solutions such as quaternion and singular value decomposition (SVD), providing transformation at iteration k with rotation and transformation ($\boldsymbol{R}_k$ and $\boldsymbol{T}_k$). Apply the transformation matrix to data, $\boldsymbol{P}_0$:
$$P = \boldsymbol{R}_k * P_0 + \boldsymbol{T}_k \tag{3}$$

3. If the change in MSE is less than a predefined threshold or the maximum number of iterations has been reached, terminate the algorithm. Otherwise start the next iteration and continue until one of the terminating conditions is reached.

ICP assumes outlier free data, which is not possible in real world applications and has no preventive measures in dealing with outliers making it sensitive to outliers.

2.2 HICP

HICP [3] considers the outlier problem by assigning weights w_i to point pairs, minimising:
$$\Sigma^2 = \sum_{i=1}^{N} w_i \|\hat{m}_i - (\boldsymbol{R}p_i + \boldsymbol{T})\|^2 \tag{4}$$

Weight values are calculated using the iterative weighted least-squares pose estimation method with outliers having a zero weighting and no influence on the registration:
$$w_i = \begin{cases} \left[1 - \frac{\|\varepsilon_i\|^2}{(cS)^2}\right]^2 & if\ \|\varepsilon_i\| \le cS \\ 0 & otherwise \end{cases} \tag{5}$$

where $\varepsilon_i = \hat{m}_i - (\boldsymbol{R}p_i + \boldsymbol{T})$ is the residual error; c is a tuning constant, and S is a scale estimator equivalent to the median of absolute distance of the entire point clouds [5].

2.3 TrICP

TrICP [4] assumes that that a minimum number of the points in the data surface can be successfully paired with a model point (minimum overlap). Any other point pairs beyond the minimum overlap are deemed as outliers and are excluded from pose estimation. TrICP computes the squared distance for each corresponding pair sorting them in ascending

order. The minimum overlap value, ξ, is found by minimising $\psi(\xi)$ in the interval [0.4, 1.0] using a Golden Selection Search where the number of paired data points $N_{po} = \xi N_p$

$$\psi(\xi) = \frac{\sum_{i=1}^{Npo} d_{i:Np}^2}{N_{p^*}\,\xi * \xi^{1+\lambda}} \tag{6}$$

The N_{po} points are used to compute the sum of the least trimmed squares distance (S'_{LTS}) of the closest N_{po} points between the data and model surface. An optimal transformation is computed using the N_{po} points minimising S'_{LTS}.

The purpose of this investigation is to establish whether ICP, TrICP and HICP are resilient to Gaussian and impulsive noise and the impact that the noise will have on the overall registration accuracy, number of iterations and the time for convergence.

3 Methodology

3.1 Image Acquisition

Four anatomical phantoms were selected for CT data acquisition, a polystyrene mannequin head; a plastic head containing tissue equivalent and real bone; a Perspex moulded foot containing real bone; and a silicon foot with 3 simulated surface ulcers (Figure 1). CT data was acquired using a Philips Brilliance 10 Slice System with a slice thickness of 1mm (www.philipsmedical.com). The scanner bed was removed from the data using the marching cubes algorithm in Analyze 10.0 (Mayo Clinic, Rochester) and 3D surfaces were extracted and stored in the stereolithography (STL) format.

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

Figure 1: The anatomical phantoms used in data acquisition (a) polystyrene head model; (b) plastic head model; (c) Perspex moulded foot; (d) silicon foot with 3 simulated ulcers.

3.2 Experimental Design

The robustness of ICP, TrICP and HICP to Gaussian and impulsive noise was investigated. For each phantom, a new set of surfaces were created with added Gaussian and impulsive noise and rotated 5^o in each orthogonal axes. All surfaces were decimated to 10% of their original facets and the noisy surfaces were registered with the original surfaces using ICP, TrICP and HICP algorithms implemented in Matlab 2009b. The addition of the Gaussian and impulsive noise is similar to the approach taken in [6].

Gaussian noise was added to the data surface with a signal-to-noise ratio of 10, 20, 30, 40 and 50 dB. Impulsive noise was added randomly to 10% and 20% of the data surfaces points. Furthermore different levels of impulsive noise were added to the data surface with the values of β tested being 0.1, 0.2, 0.3, 0.4 and 0.5 respectively.

In both investigations the termination criteria for the ICP variants were:

1. The maximum number of iterations, 300, was exceeded. This value was selected to give the algorithms sufficient iterations for convergence.

2. The change in error measure between iterations was less than $1x10^{-5}$mm.

The transformation matrix, number of iterations and convergence time were recorded. To measure the overall accuracy, the RMS between the registered non decimated noiseless surfaces was found. The RMS is the square root of the mean square distance between the closest points in the noiseless surface for every point in the noisy surface. The average convergence time, iteration and RMS were calculated.

4 Results and Discussion

4.1 Robustness to Gaussian Noise

Figure 2 (a), (b) and (c) show the impact of Gaussian noise on average RMS, time and iterations for a range of Gaussian noise levels. Results are averaged over the four phantoms with each individual phantom producing similar shape graphs, displaying similar performance. As SNR_dB decreases the level of noise increases. TrICP was the most accurate method, with RMS<0.1mm for all Gaussian levels, Figure 2(a). ICP and HICP produced similar RMS results except at 10dB where ICP had an RMS of 1.95mm and HICP had 1.76mm. From 30dB onwards all variants produced similar results <0.03mm but TrICP had a better RMS. TrICP required most iterations, even at low noise levels where the RMS with ICP and HICP is comparable, Figure 2(c). Furthermore TrICP is the slowest method except at 10dB (ICP slowest), Figure 2(b). Overall TrICP is considerably slower with HICP the fastest method for all levels of noise.

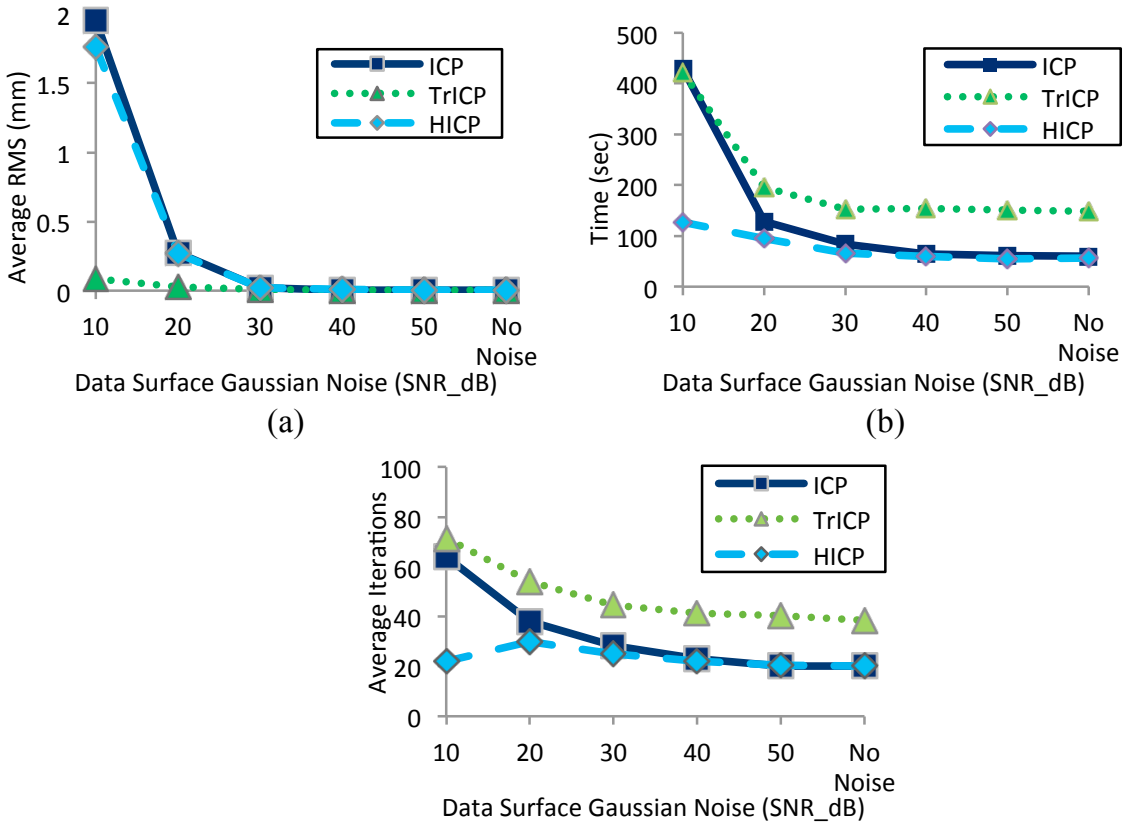


Figure 2. The effect of Gaussian noise on the average: (a) registration accuracy; (b) convergence time (sec); (c) iterations.

Although TrICP is effective in detecting outliers, this produces a high number of iterations and registration time, especially at low noise levels, showing that outlier detection is computationally expensive. Since RMS difference at low noise is negligible, consideration is required in using TrICP over HICP due to its computation requirements.

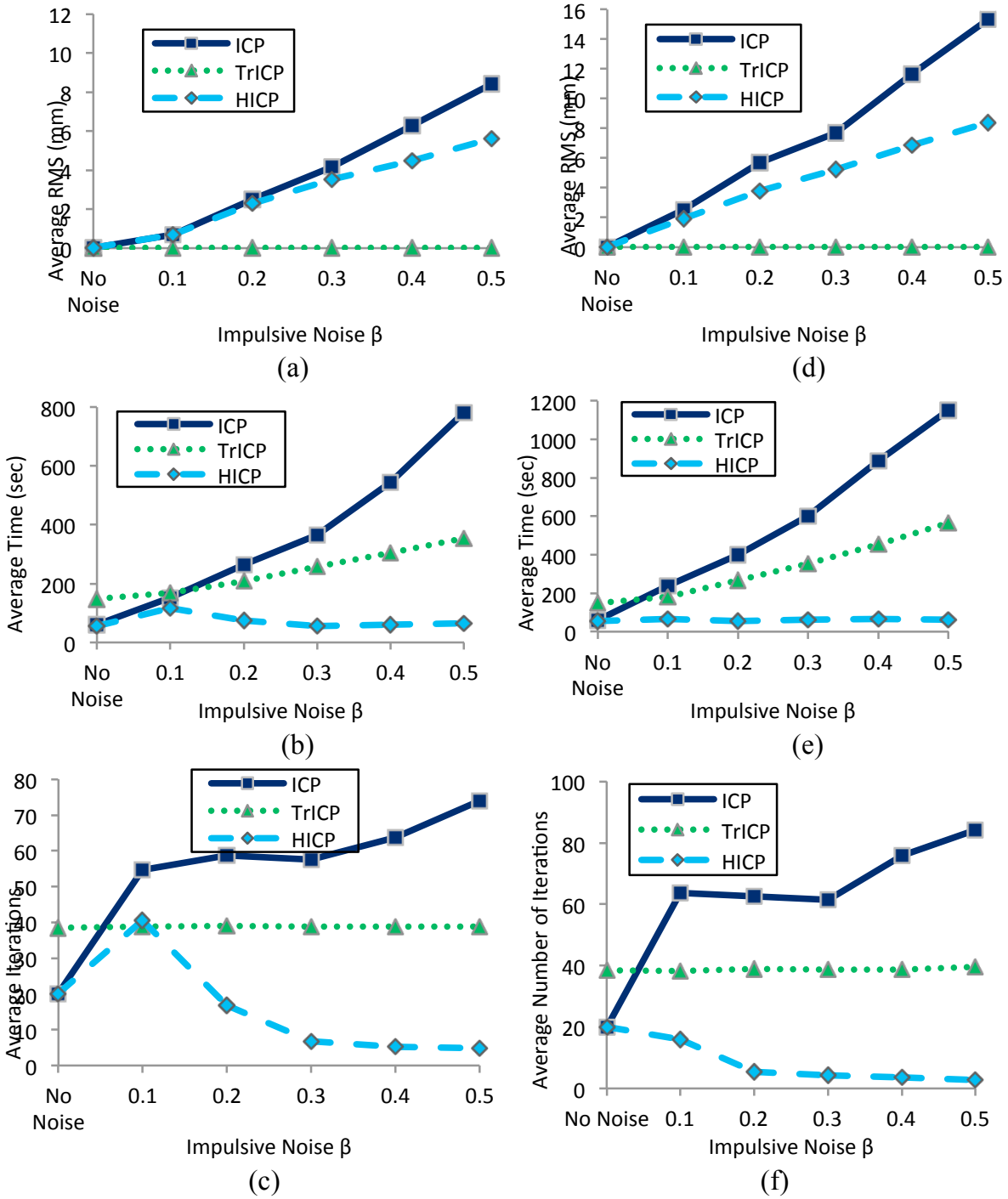


Figure 3. The effect of impulsive noise added to 10% of the data surface points on the average: (a) registration accuracy; (b) convergence time (sec); (c) iterations; and impulsive noise added to 20% of the data surface points on the average: (d) registration accuracy; (e) convergence time (sec); (f) iterations.

4.2 Robustness to Impulsive Noise

In the presence of impulsive noise in each phantom, regardless of the percentage of points changed or the noise level, TrICP produces the same RMS value showing its resilience to noise, Figure 3(a) and (d). ICP produces the largest RMS for surfaces with 10% (8.41mm) and 20% (15.32mm) of their points as outliers, showing no surface convergence. HICP performed slightly better than ICP but both methods are unsuitable for accurate registration (RMS >2mm) for impulsive noise levels of $\beta > 0.2$. TrICP requires a similar number of iterations regardless of noise, whilst HICP generally requires the fewest, decreasing as the noise level increases, indicating that HICP may get stuck in local minima with high levels of noise, Figure 3(c) and (f). ICP requires the most iterations once noise is added to the surface showing the significant impact that the noise has on the registration and the high RMS returned. HICP is the fastest of the methods, due to its small number of iterations, with ICP taking the most time when noise is added to the surface, Figure 3(b) and (e).

With impulsive noise addition TrICP is the method of choice since the RMS level does not change even with high noise and a higher percentage of points changed. Although HICP is the fastest method it is not effective in producing a suitable registration for medium to high levels of impulsive noise. At the highest impulsive noise level with 20% outliers the average HICP convergence time is 61 seconds compared with TrICP, 564 seconds, and ICP with 1152 seconds, Figure 3(e).

5 Conclusions

This investigation has shown that TrICP is robust in rejecting outliers from Gaussian and impulsive noise, but is computationally expensive so careful consideration must be given to the trade-off between required registration accuracy and the time available for registration. HICP is a viable alternative in terms of time but is only acceptable for low noise levels. ICP cannot reject outliers and is unsuccessful in registering noisy data. TrICP could be used with a manual transformation to bring both surfaces close to alignment prior to registration without the need for pre-registration data cleansing. This would remove the requirement for a landmark based registration and the issues associated with it.

References

[1] J. K. Iglehart, “The new era of medical imaging progress and pitfalls,” New England Journal of Medicine, vol. 354, no. 26, pp. 2822-2828, 2006.

[2] J. M. Phillips, R. Liu, and C. Tomasi, “Outlier Robust ICP for Minimizing Fractional RMSD,” 2006.

[3] Z. Mao, P. Sebert, and A. Ayoub, “Development of 3D measuring techniques for the analysis of facial soft tissue change,” in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2000, 2000, pp. 1051-1060.

[4] D. Chetverikov, D. Svirko, D. Stepanov, and P. Krsek, “The trimmed iterative closest point algorithm,” in International Conference on Pattern Recognition, 2002, vol. 16, pp. 545-548.

[5] R. Haralick, H. Joo, C.-N. Lee, X. Zhuang, V. G. Vaidya, and Mn. B. Kim, “Pose Estimation from Corresponding Point Data,” IEEE Transactions on Systems, Man and Cybernetics, vol. 19, no. 6, pp. 1426-1446, 1989.

[6] A. Almhdie, C. Leger, M. Deriche, and R. Ledee, “3D registration using a new implementation of the ICP algorithm based on a comprehensive lookup matrix: Application to medical imaging,” Pattern Recognition Letters, vol. 28, no. 12, pp. 1523-1533, Sep. 2007.

Correlating Histology and Spectroscopy to Differentiate Pathologies of the Colon

Duane Carey[1]
d.carey@medical-research-centre.com
Gavin Rhys Lloyd[1]
Neil Shepherd[1,2]
Nick Stone[1]
Catherine Kendall[1]
Toby Breckon[3]

[1] Biophotonics Research Unit, Gloucestershire Hospitals NHS Foundation Trust, Leadon House, Gloucester, UK

[2] Pathology Department, Gloucestershire Hospitals NHS Foundation Trust, Cheltenham, UK

[3] School of Engineering, Cranfield University, UK

Abstract

The techniques and procedures associated with histology are, in most cases, suitable for the diagnosis of colonic carcinomas. However, in cases such as epithelial misplacement the morphology of a stained tissue sample is homologous to that of cancer. This can lead to patients being misdiagnosed and undergoing unnecessary surgery.

To prevent this surgery we suggest that the epithelium of tissue samples be examined using infrared (IR) spectroscopy. In this study, IR maps of tissue sections were registered to standard histology images so that epithelial specific spectra could be collected. The differences between these spectra were explored by using Principal Component Analysis (PCA). This paper provides a novel protocol detailing how histology specific spectra can be collected. The potential usefulness of these spectra is demonstrated through the separation of epithelial misplacement cases and colonic carcinomas within PCA space.

1 Introduction

A pathologist will diagnose disease states by examining Haematoxylin and Eosin (H&E) stained tissue sections under a microscope. Staining enables the structural morphology of a tissue section to be highlighted and evaluated. Although this method of diagnosis is generally very accurate, some benign conditions can still be misdiagnosed as cancer. Epithelial Misplacement, EM, is an example of a benign pathology which often gets confused with Polyp Cancers (PC) because its morphology is homologous to that of invasive cancer.

EM polyps resemble PCs because of the environment in which they exist. EM polyps are associated with the sigmoid colon and in this region polyps are easily damaged. This is because the sigmoid colon is constantly fluxing and this movement will compress any polyps found there. The damage affects the polyps structure and forces epithelium from the exterior of the polyp into its interior. This forced movement means that when a tissue sample is sectioned for pathological assessment the sections will contain epithelial islands (indicated by a white arrow on Figure 2A). These islands are a sign indicative of invasive cancer and their presence causes pathologists to suggest that the region surrounding these benign polyps be removed from the colon [1].

However, for EM this surgery is unnecessary as only in invasive Polyp Cancers (PC) would these islands have malignant potential. In PC the islands are formed from epithelial cells which have moved under the influence of their own genetics. Therefore, the biochemistry of PC islands will be different to that of the islands within cases of EM. Infrared spectroscopy (IR) can be used to characterise these biochemical differences and facilitates the discrimination of cases of EM from PC [2], [3], [4].

2 Methods

In this study anonymous, retrospective, paraffin embedded tissue blocks were selected by a consultant histopathologist. From these blocks three contiguous 5µm thick tissue sections were cut and placed onto two normal histology slides and one IR reflective low *e* slide. One of the normal slides was stained with MNF116, a cytokeratin antibody, and the other with H&E. Images of the standard histology slides were made from the 1.25x objective lens of a Leica camera microscope. The IR images were measured using the linear array detector of the Perkin Elmer Spotlight 400 IR Spectrometer. The following protocol was then used to collect epithelial specific spectra for input into a classification model. All post collection image processing was conducted using Matlab R2007A (Mathworks, Natick USA) on a standard desktop personal computer equipped with an I7 Quad core 3.60GHz processor and 8Gb of RAM.

2.1 Histology Image generation

The standard histology images are acquired as a series of overlapping images in a grid like fashion. The overlapping images for each tissue section must first be stitched together before this sample can be further analysed. To enable this, Scale Invariant Feature Transform (SIFT) descriptors were used to find points of correspondence between the overlapping images. The SIFT algorithm finds these points from the dominant edge features of an image. It does this by constructing a difference of Gaussian pyramid for a pair of overlapping images. A subsequent search for maxima and minima within this pyramid means that keypoints [5] can be found. The minimum Euclidean distance between gradient and orientation feature vectors made for these keypoints enables points of correspondence between the overlapping images to be defined (Figure 1).

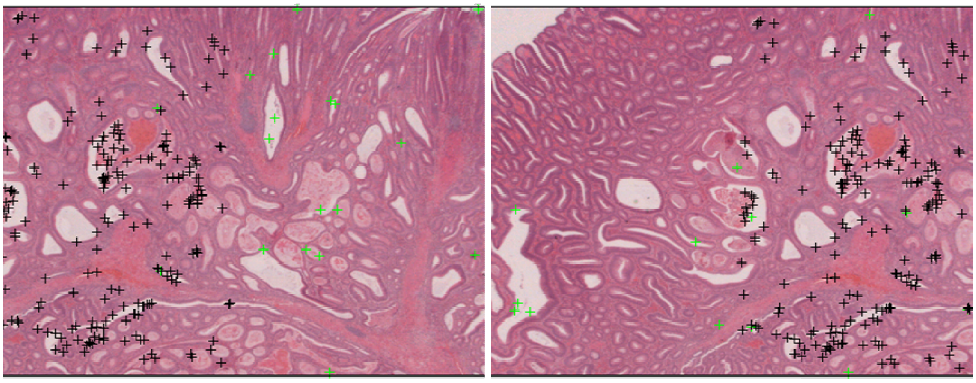


Figure 1. The matching of points of correspondence between overlapping images. The black points are the SIFT points retained after the calculation of the NCC. The green points are the SIFT points rejected.

However, the SIFT algorithm identifies many false positives but these can be removed through calculating the Normalised Correlation Coefficient (NCC) [6]. This was initiated

in this instance by defining 100 by 100 pixel segments around one of the overlapping images matched SIFT points. These image segments were used with the image it overlapped in the calculation of the NCC. As the images are of the same scene these segments should be highly correlated with regions of the overlapping image. Therefore, only SIFT points which produced a NCC above 0.95 were retained (black points in Figure 1) and allowed for the overlapping images to be stitched together. A translation was found between two images when the Euclidean distance between two consecutive SIFT points was the same [6]. This process was repeated until all of the overlapping images of a tissue section were stitched together (Figure 2A&B).

2.2 Chemical image acquisition and quality control

The IR images were acquired with a spectral resolution of 6 cm^{-1} and a spatial resolution of 25μm. Since spectra will be used within classification models it is very important that they are of an appreciable quality. To ensure this an area was summed between 1500cm^{-1} and 1600cm^{-1} wavenumbers, a biologically important area [3], [4], for all of the spectra contained within an image. Anything which was found to be two standard deviations above or below the mean of this sum was then excluded from further analysis.

2.3 Image registration

Registration of images in this work is difficult because they are acquired from different modalities. Therefore, methods which rely on the maximisation of similarity metrics are not well suited. In this instance, registration was achieved by maximising the overlap of binary masks. These binary masks were made from analysing each image of a tissue sample with PCA. The PCA scores were thresholded using a supervised *t*-test, which is equivalent to Otsu's method [7], that had some *a priori* information about the intensities of the images background pixels. This information was obtained by manually selecting a region of background pixels from a principal component that explained the variation between the background and foreground of an image. To create a binary mask using this background information a threshold is iteratively increased between the minimum and maximum of the selected principal component. At each threshold a *t*-value is produced that compares pixels below this threshold against the pre-selected background pixels. The threshold which causes the *t*-value to exceed the 95% confidence limit is used to produce a binary mask. This mask is used in the minimisation of Equation 1 so that an optimal linear transformation, T, can be found. T is composed of a rotation, translation and scaling factor and was optimised using the simplex search method [8].

$$\mathrm{E} = \sum_1^{\mathrm{N}} \sum_1^{\mathrm{M}} ((\mathrm{I2} - \mathrm{T}(\mathrm{I1})))^2 \qquad (1)$$

Here *E* is the total image registration error, *I1* represents the histology image, *I2* represents the IR chemical map and *(N, M)* represents the images spatial dimensions. All of the multimodal images were zero-padded to the same dimension.

However, these linear transformations are not sufficient because of the stretching which occurs to a tissue sample when it is being prepared. Therefore nonlinear registration is needed and this was achieved by fitting a cubic B-spline grid to an image [9]. The nonlinear transformation found was the result of moving the vertices of this grid to minimise Equation 2:-

$$E = w_i E_{img} + (w_d E_{div} + w_r E_{rot}) + w_c E_{cons} \quad (2)$$

Where *E* is the objective function to be minimised, E_{img} is a measure of the dissimilarity between the images I_2 and $T(I_1)$, $(w_d E_{div} + w_r E_{rot})$ is a regularisation term

based on the divergence and curl of the B-splines and E_{cons} is a term relating to the consistency of the registration and prevents stretching. The w_i, w_d, w_r and w_c weights allow an element of control over the deformation of the spline grid [9].

2.4 Image segmentation

The nonlinear transformations applied to the multi modal images enable very accurate feature correspondences to be achieved. This alignment allows for an image containing strong features to be used in the probing of other images where these regions of interest are less apparent. Here epithelial islands are only apparent within the standard histology images and thus to gather information on the spectroscopic characteristics of these islands the warped standard histology images were segmented. The segmentation was carried out by registering MNF116 antibody stained images and H&E images together. The MNF116 antibody is specific for a tissue sections epithelium (Figure 2 B) but it also stains other regions of a tissue sample (e.g. blood vessels). However, H&E images strongly stain blood vessels red and so the registration of the different stained images together enables the use of the Consensus Principal Component Algorithm (CPCA) algorithm [10] in preserving only the epithelium within an image. Simple manual thresholding of the CPCA super-scores allowed a binary mask to be produced and enables epithelial specific spectra to be collected and used with exploratory statistical methods.

2.5 Discriminant Analysis

In this instance, discrimination was achieved by using NIPALS PCA [11]. It facilitated the exploration of the spectroscopic variation that existed between the different pathology groups. All images of a tissue sample were vector normalised, mean centred and in the case of the IR images baseline corrected. This prevents any anomalies from affecting the results of the PCA.

The determination of significant variation by ANalysis Of VAriation (ANOVA) with 95% confidence limits assures that the different pathology groups can be separated. ANOVA uses the ratio of the intra and inter group variances against a critical value determined from the F distribution in the selection of discriminatory components. In this instance the three most discriminatory components were used so that the variation between the different pathology groups could be visualised.

3 Results

The use of histology specific spectra requires that the final image transformations be as accurate as possible. The end results are presented in Figure 2 C&D. As these results are only preliminary the root mean square error of the image differences, along with manual inspection, was taken as a measure of accuracy. The fact that these images came from different modalities means that the binary representation of the tissue foreground within these images will be slightly different. The main difference which is found is how the mucin pools are represented. This difference can be explained by the varying paraffin content that exists between the different sections produced for a tissue sample. This causes one image to have more holes within it than the other (indicated by the blue arrows on Figure 2 C&D). Even though these defects are apparent we can still determine that the epithelial islands have been accurately registered. This is because in the interior of the

difference image the only significant change apparent is around the edges of the images epithelial islands (indicated by red arrows on Figure 2 C&D).

From these registered images epithelium specific spectra was collected via image segmentation. In this case the output of the CPCA algorithm (Figure 2E) which acted on images presented in Figure 2 A&B was used to generate epithelial specific spectra.

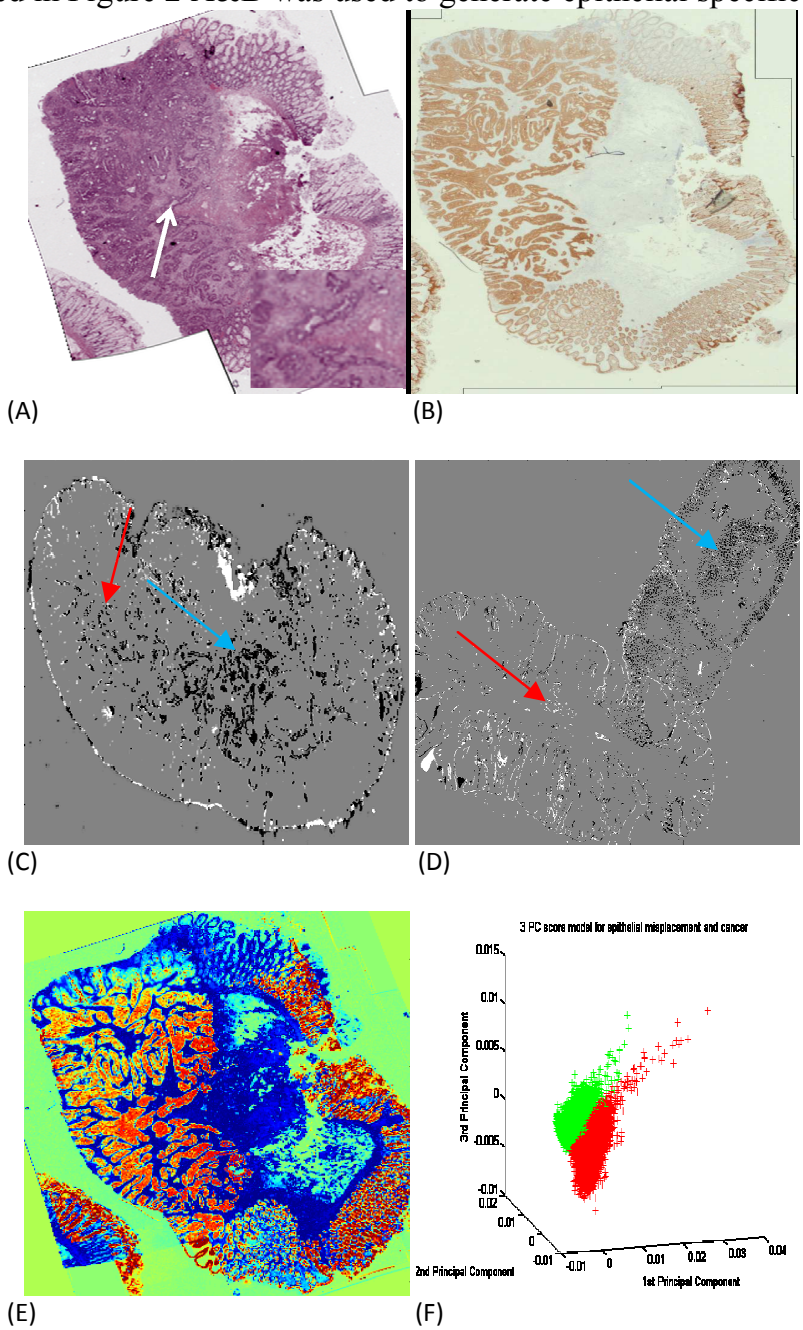


Figure 2. A depiction of how epithelium specific spectra can be collected and used: (A & B)Unregistered MNF116 and H&E histology images stitched together from images acquired from the Leica camera microscope (the white arrow indicates an epithelial island which has been enlarged in the bottom of A); (C & D)the differences between the multimodal images after application of an nonlinear transformation (blue and red arrows indicate the observer criteria used for accuracy);(E)Result of the CPCA algorithm acting on combined MNF116 and H&E images; (F)3D PCA score plot made from the epithelial specific spectra (red points are PC spectra and the green points are EM spectra).

This method allowed for 9,804 spectra to be collected from six samples, three from each pathology group and roughly half of the spectra were contributed from each group. These spectra were used within a PCA model so that the variation between the different pathology groups could be explored. This exploration proved fruitful as is evident from

Figure 2F and potentially confirms that epithelial specific spectra, the green points in Figure 2F, can indeed be used to differentiate epithelial misplacement from cancer, the red points in Figure 2F.

4 Conclusion

A procedure for the intermodal registration of digital histology images and IR spectroscopic images is described and successfully applied to 6 samples. An image segmentation algorithm was then applied to the registered images so that epithelial islands could be located and their associated spectra collected. PCA was then used to show the potential of these abstracted spectra in discriminating EM from PC (Figure 2F). This type of approach could lead to improved cancer diagnostics and reduce the number of EM patients receiving unnecessary treatments [12].

The enlargement of the sample size is an important aspect of future work along with automating the classification. The automated discrimination of EM from cancer can be achieved with Linear Discriminant Analysis (LDA) or Support Vector Machine (SVM) [13] techniques and will help validate the PCA model presented in Figure 2F. Another important future aspect to consider is the method of error analysis used to determine the accuracy of registration. As the collection of specific spectra rests on the accurate registration of images more advanced consistency methods between the modalities of a tissue sample will be evaluated in the future [14].

[1] V. C. Petersen, a L. Sheehan, R. L. Bryan, C. P. Armstrong, and N. A Shepherd, "Misplacement of dysplastic epithelium in Peutz-Jeghers Polyps: the ultimate diagnostic pitfall?," *The American journal of surgical pathology*, vol. 24, no. 1. pp. 34-9, Jan-2000.

[2] C. Kendall et al., "Vibrational spectroscopy: a clinical tool for cancer diagnostics.," *The Analyst*, vol. 134, no. 6, pp. 1029-45, Jun. 2009.

[3] J. T. Kwak, S. M. Hewitt, S. Sinha, and R. Bhargava, "Multimodal microscopy for automated histologic analysis of prostate cancer.," *BMC cancer*, vol. 11, no. 1, p. 62, Jan. 2011.

[4] J. J. Wood, C. Kendall, G. R. LLoyd, N. A Shepherd, T. A Cook, and N. Stone, "Infrared spectroscopy to estimate the gross biochemistry associated with different colorectal pathologies," *Imaging*, vol. 8087, p. 80870P-80870P-9, 2011.

[5] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*. vol. 60, no. 2, pp. 91-110,2004.

[6] V. Rankov, "An Algorithm for image stitching and blending," *Proceedings of SPIE*, vol. 5701, pp. 190-199, 2005.

[7] J.-hao Xue and M. D. Titterington, "t-Tests, F-Tests and Otsu's Method for Image Thresholding," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2392-2396, 2011.

[8] J.C .Lagarias, J. A. Reeds, M. H. Wright and P. E. Wright, "Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions," *SIAM Journal of Optimization*, vol. 9, no 1, pp. 112-147, 1998.

[9] I. Arganda-carreras, C. O. S. Sorzano, and R. Marabini, "Consistent and Elastic Registration of Histological Sections using Vector-Spline Regularization," *Computer Vision Approaches to Medical Image Analysis In Computer Vision Approaches to Medical Image Analysis*, vol. 4241, pp. 85-95, 2006.

[10] J. A. Westerhuis, T. Kourti, and J. F. Macgregor, "Analysis of multiblock and hierarchical pca and pls models," *Journal of Chemometrics*, vol. 321, pp. 301-321, 1998.

[11] R. Brereton, "Chemometrics. Data Analysis for the Laboratory and the Chemical Plant," *Wiley,* pp 194, 2004

[12] F. L. Greene, "Epithelial Misplacement in adenomatous polyps of the colon and rectum," *Cancer*, vol. 33, no. 1, pp. 206-217, Jan. 1973.

[13] J.W. Han, T.P. Breckon, D.A. Randell, and G. Landini, "The Application of Support Vector Machine Classification to Detect Cell Nuclei for Automated Microscopy", *In Machine Vision and Applications*, Springer, vol 23, no. 1, pp. 15-24, 2012.

[14] E. T. Bender and W. A. Tome, "Utilization of consistency metrics for error analysis in deformable image registration," *Medical Physics*, vol. 54, no. 18, pp. 5561-5577, 2009.

Performance evaluation of simultaneous RGB analysis for feature detection and tracking in endoscopic images

F. Selka[12]
selka.faical@gmail.com
S. Nicolau[1]
stephane.nicolau@ircad.u-strasbg.fr
A. Bessaid[2]
a.bessaid@gmail.com
L. Soler, J. Marescaux[1]

[1] IRCAD Strasbourg, France
[2] GBM
Abou Bekr Belkaid University
Tlemcen, Algeria

Abstract

In laparoscopic surgery, soft tissue motion tracking and 3D structure reconstruction are crucial to provide an augmented reality view in navigation systems. Performing an accurate real-time surface 3D reconstruction requires an efficient detection of interest points. In this paper, we propose an approach to increase the number of good features to track and to improve their tracking robustness in endoscopic images based on simultaneous RGB analysis, instead of gray level only. The proposed method has been evaluated on human and pig endoscopic images, using Shi-Tomasi, SURF and SIFT feature detector and Lucas-Kanade tracking algorithm. Results confirm that our approach increases the number of detected features up to 40% and avoids wrong tracking of about 17% of points, in comparison with gray level channel.

1 Introduction

Minimally invasive surgical (MIS) procedures are gaining popularity in the medical community for their ability to reduce patient recovery time, patient morbidity and patient trauma. The main tool used by surgeons is an endoscopic camera, which is inserted through an orifice (natural or artificial) into the human body. However, this technique has also drawbacks for surgeons such as field of view limitation and 3D vision loss, which lengthen intervention duration. These limitations encouraged several areas of research in the field of the computer vision. Most of these researches try to provide real-time information to surgeons to decrease the limitations mentioned above [9]. Recovering in real-time the 3D geometry of the abdominal cavity could, for example, allow piloting automated systems, with the aim of assisting surgeons and increasing intervention safety. These systems are usually based on real-time organ surface reconstruction, which relies on tracking and matching of feature points. The surface reconstruction quality thus highly depends on this first tracking step.
In this paper, we highlight that the standard approaches for feature tracking in endoscopic

images are performed on gray level images, and show that considering all RGB information increases the number of good features and improves the robustness of the tracking step in endoscopic images.

Generally, methods proposed for 3D surface reconstruction in MIS [9, 10, 12], are based on tracking regions of interest in successive image sequences. The problem of locating a region of interest in one image and finding the corresponding region in another one is difficult in MIS since images can be low in contrast, noisy and poorly illuminated[9]. To solve this problem, an algorithm must be used to identify the most robust points for the preliminary step. Feature point detectors based on cornerness measures like Harris[5] and Shi-Tomasi[11] are still famous and used to extract points[10]. More recently, descriptors like SIFT[8], SURF[3], have been widely used to describe and match a region of interest. Both approaches were integrated for tracking deformable soft tissues in MIS[10, 12]. To reduce algorithm complexity and time computation, most works in endoscopic surgery use gray level images to identify features instead of using the full RGB information. Mountney et.al.[10] suggest that color does not seem to bring significant improvement. However, no quantitative evaluation has been provided. Although we agree that computational time is important, it seems also important to provide as much robust and spread points as possible. Point robustness is important to avoid relying too much on a supplementary detection step of wrong matches using RANSAC [4] or removal outliers based on epipolar constraints [6], which can be computationally expensive. In this paper, we show that a simulatenous analysis of RGB channels allows for better identification and robustness of features in endoscopic images. Our approach has been motivated by recent works related to narrow (NBI) or multi (MBI) band imaging techniques [7]: using a specific linear combination of R, G, B channels can highlight the visualisation of different tissues which do not absorb identical wavelengths. The remaining part of this paper is structured as follow: In section 2, we firstly argue that considering simultaneous RGB analysis allows to increase the number of robust features compared to the use of gray-level image only. Then, we explain how we use the 3 RGB channels to identify and track features. In section 3, we show on human and pig *in-vivo* data that our approach increases the number of good features up to 40% and that we track more points than in gray level channel with a better robustness.

2 Simultaneous detection and tracking on RGB channels

The use of the RGB space is very common in image processing since it is provided by most acquisition devices. In RGB space, each signal corresponds to a different wavelength band of the visible spectrum. The gray level (BW) image is a linear combination of the 3 signals: $BW = 0.299 \times R + 0.587 \times G + 0.114 \times B$.

These weights depend on the exact choice of the RGB primaries, the ones we provide in previous equation are typical [1]. Usually, good features are selected using a detector based on gradient intensity. A basic analysis highlights the 2 following drawbacks. On the one hand, if one point is detected on B channel for instance, and not on the others, it is then likely that this point will not be detected in BW image, since the gradient intensity in BW will be weighted on all channels and B contribution is too low. Thus, a point that could be a good feature to track in an independent channel will not be tracked. On the another hand, a point detected in BW image may be tracked more efficiently in the channel in which its gradient properties are stronger. For these reasons, we propose to perform a simultaneous analysis of RGB channels. Our method can be divided in 2 main steps, the selection of good

features to track in R, G, B channels, and the tracking strategy.

2.1 Selection of features to track

In this subsection we firstly describe how we select the features that will be tracked along the endoscopic video sequence. We choose to detect features using Shi-Tomasi algorithm [11]. Let S_x be the feature response of the channel x : S_x is a set of 2D point coordinates. At the very beginning of the video i.e. frame 0, we firstly compute S_R, S_G and S_B and consider the union of these point sets $S_R \cup S_G \cup S_B$. Secondly, we merge points in $S_R \cup S_G \cup S_B$ which are very close. Typically, a point $M(x,y)$ and a point $U(u,v)$ are considered identical if $(x+\tau > u > x-\tau)$ and $(y+\tau > v > y-\tau)$, and the new point is the average of M and U. τ is chosen considering the resolution of the image and the size of the observed scene. In our case, we use a HD camera and the scene has a rough size of 30 cm, it is then reasonable to choose $\tau = 5$ pixels since it corresponds to 0.7 mm. Finally, the new set after the point merging process is called S^0_{all}. Note that the points in S^0_{all} are no longer associated to a specific color channel. In fact, we consider that all of them can be a good feature to track in all channels. This choice may seem odd, but we have experimentally noticed that due to illumination change a good feature in R channel only, for instance, can become a good feature in B or/and G channels after several frames.

Points in S^0_{all} will be tracked along frames, the updated point set in frame i will be denoted S^i_{all}. After many frames, several points in S^i_{all} are no more visible in the video sequence due to the camera movement. Once 10 points have moved out of the endoscopic image, we decide to launch again Shi-Tomasi detector on R, G and B channels, to compute $S_R \cup S_G \cup S_B \cup S^i_{all}$ and to merge points which are too close, obtaining an updated S^i_{all} which now includes new points from the simultaneous detection on RGB channels. This process is performed each time that 10 points have left the endoscope field of view to reduce computational local. The next subsection explains how points in S^{i+1}_{all} are estimated from their position in the previous frame S^i_{all}.

2.2 Tracking strategy

Let P^i be a point in S^i_{all} in frame i. In this subsection we explain how we compute the position P^{i+1} in frame $i+1$ from its estimated position P^i in the previous frame. Firstly, we estimate on R, G and B channels the motion of P^i in frame $i+1$ using Lucas Kanade algorithm [2]: $P^{i+1}_R, P^{i+1}_G, P^{i+1}_B$. Ideally, if the tracking was perfect we should observe $P^{i+1}_R \simeq P^{i+1}_G \simeq P^{i+1}_B$. Moreover, since we assume that the endoscope motion varies slowly, the motion of P^i (i.e: $\|P^i\ P^{i+1}\|$) should have a magnitude close to the average motion of $P^{i-2}, P^{i-3}, ..., P^{i-n}$ on the previous frames. Mathematically, this means that the distance m^{i+1} of P^{i+1}_R, P^{i+1}_G and P^{i+1}_B to their gravity center O^{i+1} : $m^{i+1} = \frac{O^{i+1}P^{i+1}_R + O^{i+1}P^{i+1}_G + O^{i+1}P^{i+1}_B}{3}$ should be extremely small (cf. Fig 1 a). Practically, $P^{i+1}_R, P^{i+1}_G, P^{i+1}_B$ are not identical, and their motions can be inconsistent.

In the best case, the distance between these 3 estimations is very small and motion estimation is consistent, it is then likely that they are all reasonable estimations and we then decide that our estimation of P^{i+1} in frame $i+1$ will be equal to $\frac{P^{i+1}_R + P^{i+1}_G + P^{i+1}_B}{3}$. More generally, the computation of P^{i+1} will depend on the scattering of $P^{i+1}_R, P^{i+1}_G, P^{i+1}_B$ (represented by m^{i+1}) and on their motion $d^{i+1}_R = \|P^{i+1}_R P^i\|$, d^{i+1}_G and d^{i+1}_B.

Briefly, we discard in a first step P^{i+1}_x if d^{i+1}_x is above a threshold $\lambda_{\overline{d}}$. In a second step, we compute the scattering magnitude m^{i+1} with the remaining points and estimate P^{i+1} depending on m^{i+1}. Fig.1 a-f describes all possible cases, the yellow circle representing the

threshold of tolerated scattering $\lambda_{\overline{m}}$. Note that Lucas Kanade algorithm sometimes fails at providing an estimation (in that case, the feature is called *lost* point in Sec. 3.2). Practically, it can happen that P^i motion in frame $i+1$ is not estimated on channel x. In our algorithm, this case is identical to $d_x^{i+1} > \lambda_{\overline{d}}$.

The two thresholds $\lambda_{\overline{d}}$, $\lambda_{\overline{m}}$ are defined as follow. Let $\overline{d}$ be the average displacement of several points P selected in ROI located in the image center over the last 20 frames $\overline{d} = \sum_{j=0}^{20} \frac{\|P^{i-j}\ P^{i+1-j)}\|}{20}$ and $\sigma_{\overline{d}}$ the standard deviation of these distances. We suppose there is no tracking failure over these 20 frames since the illumination condition does not change in the image center. In a similar way, $\overline{m}$ is the average of m^i on the 20 previous frames, and $\sigma_{\overline{m}}$ its standard deviation. We choose $\lambda_{\overline{d}} = \overline{d} + 3\sigma_{\overline{d}}$ and $\lambda_{\overline{m}} = \overline{m} + 3\sigma_{\overline{m}}$. $\overline{d}$, $\overline{m}$ and their standard deviation will be re-estimated every 20 frames.

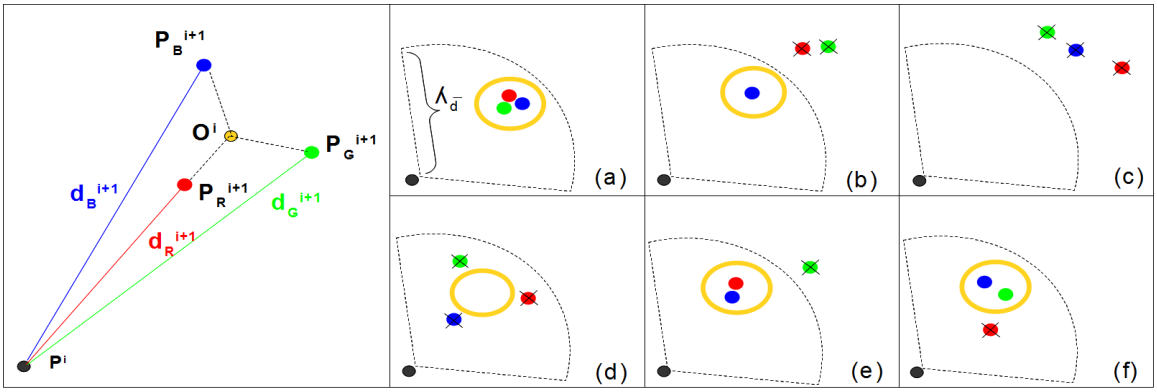


Figure 1: Left: $P_R^{i+1}, P_G^{i+1}, P_B^{i+1}$ are the estimated position of P^i in frame i+1, and O^{i+1} their gravity center. Right: these 6 figures sketch possible tracking scenario during simultaneous tracking in RGB channels. In each case, we firstly discard estimation in channel x if d_x^{i+1} is above $\lambda_{\overline{d}}$ (case b,c,e). Then, if m_i is below $\lambda_{\overline{m}}$, P^{i+1} is the average of the remaining points (case a,b,e). If $m_i > \lambda_{\overline{m}}$ (case d,f), we check if the distance between 2 points is below $\lambda_{\overline{m}}$ and average them (case f). If not, the tracking is stopped (case d).

3 In-vivo evaluation of RGB simultaneous analysis

We evaluate our approach on 4 *in-vivo* HD video sequences corresponding to an abdominal exploration. Two of them contain a human liver and gallbladder and the other two contain pig bowels (cf Fig.2). From left to right, the sequence contains respectively 550, 675, 625, 875 frames. We used a 3 CCD HD endoscope Storz Image1 Hub with resolution of 1920×1080 and 0° degree endoscopic lens. In this evaluation, we firstly evaluate the supplementary amount of features using all RGB channels with Shi-Tomasi, SURF and SIFT detectors, in comparison with gray level only. Secondly, we show that the feature set is more robust when tracked with our approach than in gray level channel only.

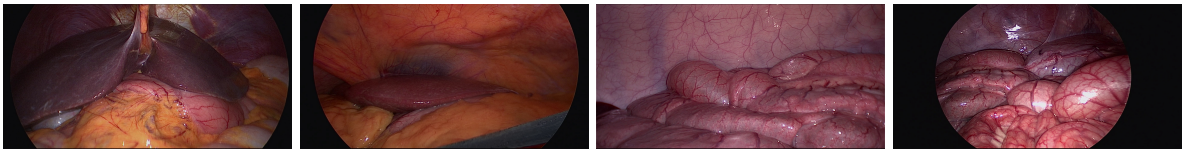


Figure 2: Image sample of the different sequences used in our analysis.

3.1 RGB simultaneous analysis to increase the amount of features

We evaluate our feature selection method on all frames of the 4 endoscope sequences using Shi-Tomasi, SIFT and SURF algorithm and compare it to feature selection on gray level

image only. The Opencv library was used with the following parameters, which were used on all sequences: for Shi-Tomasi the threshold was defined as 0.02 with minimum distance of 35. The derivatives were calculated using generalized Sobel with aperture size = 3. A smoothed window of a Gaussian ($\sigma = 1$) with (3×3) size was used to average the derivatives. For SIFT: octave = 4, threshold 0.05 and edge threshold = 5 was taken. For SURF: hessian threshold = 800. Tab.1 shows the average percentage of new features compared to features detected in gray level image for the liver and the bowel sequences. One can see that considering simultaneous detection in all channels can provide more than 40% new features for Shi-Tomasi and SURF and more than 30% for SIFT. Fig.3 provides an example of our feature detection method. It is worthy to note that 99% of points detected in BW image are also selected with our approach.

Detector	Shi-Tomasi		SIFT		SURF	
Sequence	Bowels	Liver	Bowels	Liver	Bowels	Liver
Average	46.63%	44.62%	34.84%	36.50%	42.27%	40.45%

Table 1: Average percentage of new features using simultaneous RGB analysis, compared to gray level analysis only.

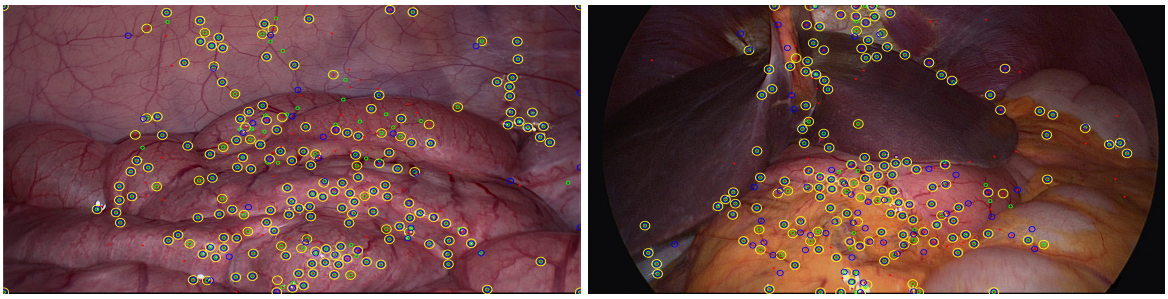


Figure 3: Example of feature detection using Shi-Tomasi. Red/green/blue/yellow circles correspond to features detected in R/G/B/BW channels.

3.2 Evaluation of feature robustness using simultaneous RGB tracking

To compare the tracking robustness of our method with the standard tracking based on gray level image, we propose the following. We select features in each video sequence using our approach. Then, we track all points along sequences and visually report features, which are *lost* by Lucas-Kanade algorithm or *wrong* (tracked feature does no longer represent the original feature). We also report the percentage of frames during which the feature is *lost* or *wrong*. The Lucas Kanade algorithm was used with a (25×25) window size for block matching and a pyramid level set to 3. Tab.2 provides the average results for the feature set tracked on *BW* image and using our method (*RGB*). One can see that our approach allows to avoid losing almost all points, which were *lost* in *BW* image. This means that each tracked feature was always found in at least one of the 3 RGB channels. Results also show that all selected features have been consistently tracked along all sequences, whereas we report an average of about 7 wrong points tracked in *BW* image. On average, our approach increases up to 17% the number of points which are properly tracked.

Liver	BW	RGB	Bowels	BW	RGB	Liver	BW	RGB	Bowels	BW	RGB
Lost point	33	0	Lost point	52	0	Wrong point	8	1	Wrong point	7	0
Lost frame	0.97%	0%	Lost frame	1.67%	0%	Wrong frame	1.07%	0.2%	Wrong frame	3.06%	0%

Table 2: The average result for the two sequences. liver/(bowel). tracking on 302 points over 404 frames /(tracking on 282 points over 360 frames.)

4 Conclusion

This paper presents a method for feature detection and tracking in endoscopic images based on simultaneous RGB analysis, which allows to provide more features and robustness than using gray-level images. This method was evaluated on patient and pig data and results show that up to 40% (resp. 40%, 30%) of supplementary points can be detected using Shi-Tomasi method (resp. SURF, SIFT descriptor). We also show that the proposed method for feature tracking increases the number of robust points up to 17%. We believe this work has highlighted that using all RGB information for detection and tracking of features can nicely decrease the number of outliers for the usual next step: shape reconstruction from motion. In the future, we will strengthen our evaluation on more organs (stomach, pancreas, kidney). We also plan to investigate several linear and non-linear combinations of RGB to enhance specific tissues or organs using wavelength response of each structure.

References

[1] OpenCV. In *http://opencv.willowgarage.com/documentation*, 2010.

[2] S. Baker and I. Matthews. Lucas-kanade 20 years on: A unifying framework. *International Journal of Computer Vision*, 56(3):221–255, 2004.

[3] H. Bay, T. Tuytelaars, and L. Van Gool. Surf: Speeded up robust features. *Computer Vision–ECCV 2006*, pages 404–417, 2006.

[4] O.G. Grasa, J. Civera, and JMM Montiel. Ekf monocular slam with relocalization for laparoscopic sequences. In *ICRA 2011*, pages 4816–4821. IEEE.

[5] Chris Harris and Mike Stephens. A combined corner and edge detector. In *In Proceedings of The Fourth Alvey Vision Conference (1988)*, pages 147–152, 1988.

[6] M. Hu et. al. Reconstruction of a 3d surface from video that is robust to missing data and outliers: Application to minimally invasive surgery using stereo and mono endoscopes. *Medical Image Analysis*, 2010.

[7] S. Kodashima and M. Fujishiro. Novel image-enhanced endoscopy with i-scan technology. *World Journal of Gastroenterology: WJG*, 16(9):1043, 2010.

[8] D.G. Lowe. Object recognition from local scale-invariant features. In *ICV 1999*, volume 2, pages 1150–1157. Ieee.

[9] P. Mountney, D. Stoyanov, and G.Z. Yang. Three-dimensional tissue deformation recovery and tracking. *Signal Processing Magazine, IEEE*, pages 14–24, 2010.

[10] P. Mountney et. al. A probabilistic framework for tracking deformable in minimally invasive surgery. *MICCAI 2007*, pages 34–41.

[11] J. Shi and C. Tomasi. Good features to track. In *Proceedings CVPR'94*, pages 593–600. IEEE, 1994.

[12] D. Stoyanov et. al. Soft-tissue motion tracking and structure estimation for robotic assisted mis procedures. *MICCAI 2005*, pages 139–146, 2005.

Scan-rescan reproducibility of neurite microstructure estimates using NODDI

Maira Tariq[1]
maira.tariq.11@ucl.ac.uk
Torben Schneider[2]
t.schneider@ucl.ac.uk
Daniel C. Alexander[1]
d.alexander@cs.ucl.ac.uk
Claudia A.M. Wheeler-Kingshott[2]
c.wheeler-kingshott@ucl.ac.uk
Hui Zhang[1]
g.zhang@cs.ucl.ac.uk

[1] Department of Computer Science &
Centre for Medical Image Computing
University College London
London, United Kingdom

[2] NMR Research Unit, Department of
Neuro-inflammation
UCL Institute of Neurology
University College London
London, United Kingdom

Abstract

In this work we provide a preliminary assessment of the reproducibility of the Neurite Orientation Dispersion and Density Imaging (NODDI), a recent diffusion MRI technique for directly quantifying microstructural indices of neurites *in vivo*, in the human brain. It is important to assess the reproducibility of such a technique to verify the precision of the method, which has implications for translation to clinical studies. NODDI outputs indices which reflect the functional and computational complexity of various regions of the brain and thus can provide useful information, non-invasively, for understanding pathology of the brain. We compare the parameter maps derived from diffusion MRI data acquired using the NODDI protocol from a normal subject, at two separate imaging sessions. We show that the NODDI indices have reproducibility comparable to that of the DTI indices. We additionally show that the clinically feasible NODDI protocol maintains good reproducibility of parameter estimates, comparable to that of a more comprehensive protocol.

1 Introduction

NODDI is a practical diffusion MRI technique, introduced in [15], for estimating the integrity of the axons and dendrites, collectively known as neurites, in the human brain. Dendrites and axons are the extensions from the neural cell bodies, which have a vital role in the communication network of the brain. The morphology and complexity of these structures have been shown to indicate the function as well as pathology of the brain[8, 9, 10]. For example dispersion of dendrites is linked with brain development [6], and their density with ageing [9]. Thus by quantifying the physical characteristics of neurites, promising markers of progression of brain diseases can be obtained and used in the diagnosis, prognosis and treatment of brain disorders.

Neurite density and their *orientation dispersion* are the two microstructural parameters estimated by NODDI. This is in contrast to Diffusion Tensor Imaging (DTI) [4], the standard clinical diffusion MRI technique, which outputs Fractional Anisotropy (FA) and Mean Diffusivity (MD). Although both NODDI and DTI probe microstructure from the diffusion of the water molecules within the brain, DTI is limited in its ability to provide information about the specific changes in microstructure [12]; the indices obtained are affected simultaneously by a number of microstructural changes (e.g. demyelination, inflammation, axonal loss, gliosis), which give rise to the same alterations of their values. NODDI on the other hand estimates the microstructure directly, using an analytical model relating these parameters to the diffusion MRI signal. The NODDI parameters have been shown in [15] to disentangle the microstructural indices which contribute to the changes in FA and thus can provide information about specific changes in microstructure for particular pathologies.

NODDI is underpinned by a multi-compartment signal model of tissue diffusion, which describes analytically how presence of certain structures within the brain affects the diffusion MR signal. Numerous similar techniques exist to infer microstructural features directly (See [3] for a review), but differ from NODDI in the features they model (mostly model only white matter), as well as how they are modelled. But the main limitation of the other models is that they are not clinically feasible. The original work [15] shows results which validate and evaluate the accuracy as well as clinical viability of NODDI, but the precision of the parameters estimated, over multiple imaging sessions, is not explored.

In this work we evaluate the precision of the NODDI by assessing its reproducibility for estimating the microstructural parameters. Such an assessment is important for justifying the use of this technique in clinical and research studies. This is especially true for longitudinal studies since the variability in the parameters needs to be purely due to the changes in brain microstructure.

The paper is organised as follows: Section 2 details the tissue model adapted by NODDI, the data acquisition and the fitting procedure; Section 3 contains the results; Section 4 evaluates and summarises the key findings and discusses future work.

2 Materials and Methods

NODDI combines a three-compartment tissue model with a two-shell high-angular-resolution diffusion imaging (HARDI) protocol, optimised for clinical feasibility. The following sections summarise the NODDI model used to infer the microstructure indices, as well as the protocols used to acquire the images.

2.1 NODDI tissue Model

NODDI is an analytical tissue model, which enables differentiation of three microstructural environments: Intracellular (IC), Extracellular (EC) and CerebroSpinal Fluid (CSF). Each environment affects the diffusion of water molecules within and around them in a characteristic manner and thus contributes distinguishingly to the normalised MR signal, obtained from diffusion imaging. The full normalised signal in a NODDI model is formed from contributions from all of these environments, as shown in the following equation:

$$A = (1 - v_{iso})[v_{ic}A_{ic} + (1 - v_{ic})A_{ec}] + v_{iso}A_{iso} \quad (1)$$

The *intracellular* compartment, representing the neurites, has a neurite density given by the volume fraction v_{ic}, and contributing diffusion signal A_{ic}. Axons and dendrites are structurally similar to cylinders, and are modelled for simplicity as sticks, i.e. zero radius cylinders. This reflects the highly restricted diffusion perpendicular to and unhindered diffusion along the length of the structures. The IC diffusion signal also incorporates the orientation dispersion of the neurites, utilising a *Watson Distribution* [11]. Watson distribution is described by κ, which measures the extent of orientation dispersion about μ, the dominant orientation. The *orientation dispersion index* (ODI), output by NODDI, is defined as in equation (2) and ranges from 0, for coherently oriented structures, to 1 for isotropic structures.

$$OD = (2/\pi)\arctan(1/\kappa) \tag{2}$$

The glial cells and, in grey matter, the cell bodies make up the *extracellular* compartment of NODDI. Diffusion of molecules in this compartment is hindered by the presence of neurites, and is thus modelled by Gaussian anisotropic diffusion, i.e. a cylindrically symmetric tensor.

The final compartment models the CSF as isotropic Gaussian diffusion.

2.2 Data Acquisition

Diffusion MR data was acquired on a Philips 3T system with maximum gradient strength of 60mT/m. One healthy volunteer was scanned twice over two sessions separated by 4 weeks. Same imaging protocol was used as the original study [15], which was optimised using the technique described in [1], with a constraint of 30 mins on acquisition time and maximum gradient strength of 60 mT/m. This optimised NODDI protocol consists of a $711 s/mm^2$ shell with 30 gradient directions and a $2855 s/mm^2$ shell with 60 directions, and 9 b=0 measurements. The NODDI protocol is clinically feasible as it takes less than 30 mins to acquire. A more comprehensive 4-shell protocol was utilised to provide a *pseudo-gold-standard* to assess the reproducibility of the NODDI protocol against and determine the effect of acquiring fewer shells on the reproducibility of the parameters estimated. An additional $1000 s/mm^2$ shell of 30 directions and a $2000 s/mm^2$ shell of 60 directions were thus acquired and combined with the NODDI protocol to form this comprehensive protocol. The results from the two protocols were compared to assess the precision in estimation of the indices, when utilising more clinically feasible protocol.

2.3 Processing

The model was fitted to the acquired data using the procedure described in [2], without the MCMC procedure, for computational efficiency [15]. The parameter maps obtained were subsequently aligned using DTI-TK [13] to enable voxel-wise assessment of parameter reproducibility. To assess the parameters' reliably, the maps were segmented into White Matter (WM), Grey Matter (GM) and CSF regions, using MD and *linearity measure* [14] of the diffusion tensor. CSF was segmented as areas where MD was higher than 80%, WM as linearity higher than 0.2 and GM as rest of the brain regions.

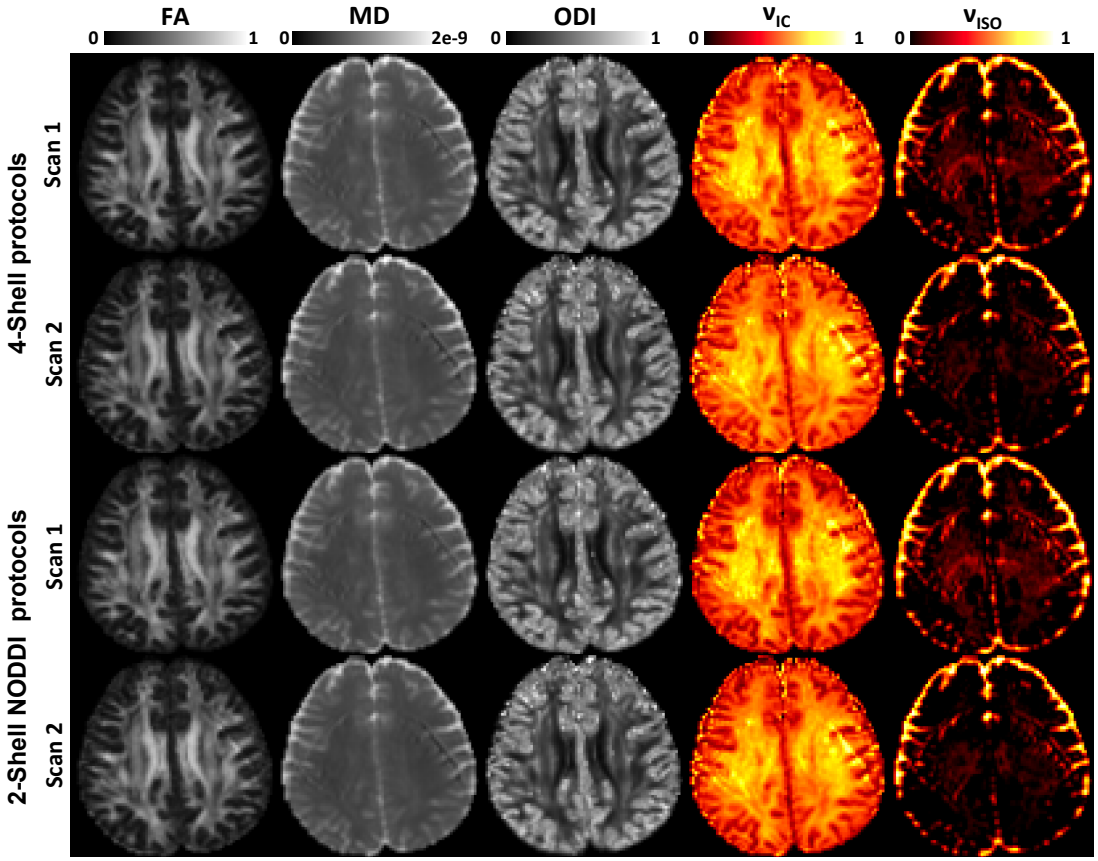


Figure 1: Scan-rescan parameter maps obtained for mid-axial slice, for DTI (first two columns) and NODDI (last three columns), using 4-Shell (first two rows) and NODDI (last two rows) protocols

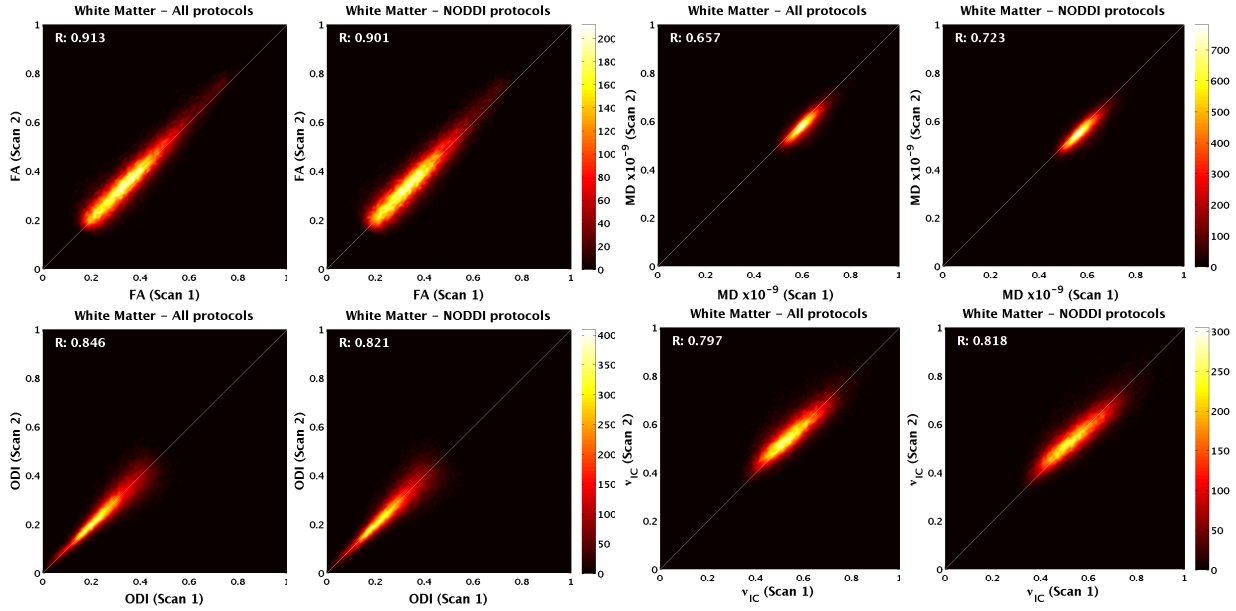


Figure 2: Scatter plots of scan-rescan for WM voxels, for DTI indices (top row) and NODDI parameters (bottom row). R values shown in plots represent the correlation coefficient for the two scans, for each parameter

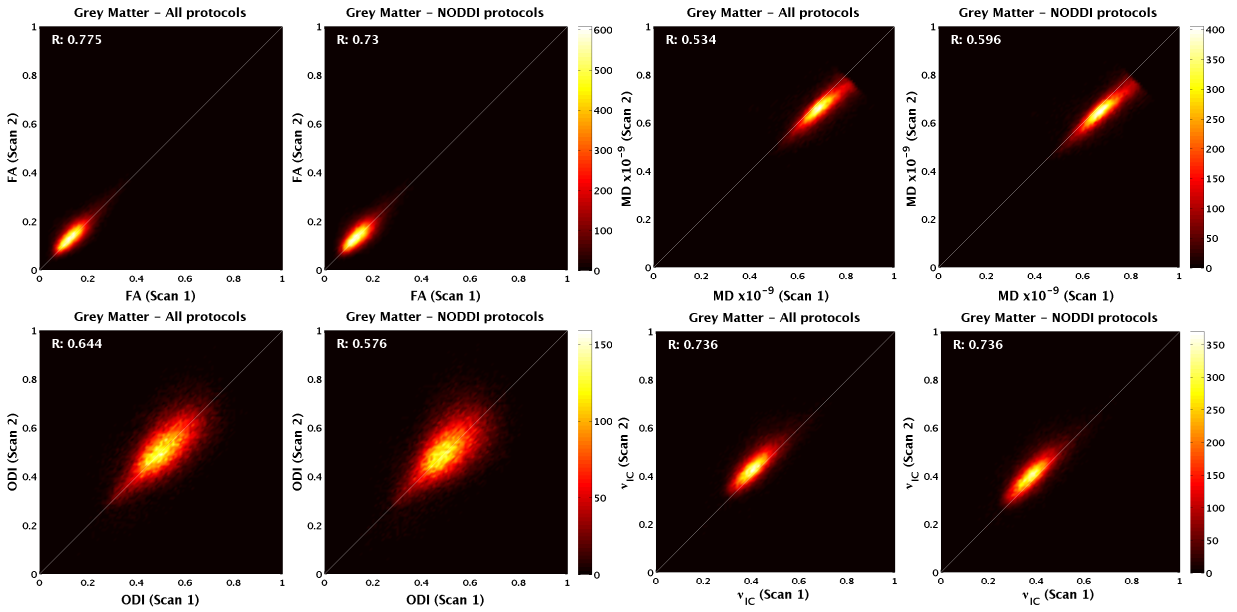


Figure 3: Scatter plots of scan-rescan for GM voxels, for DTI indices (top row) and NODDI parameters (bottom row). R values shown in plots represent the correlation coefficient for the two scans, for each parameter

3 Results

Figure 1, using an exemplar slice, illustrates qualitatively the scan-rescan reproducibility of the NODDI parameters in comparison to the standard DTI indices. The scatter plots in Figures 2 and 3 demonstrate this over the whole brain and separately for WM and GM regions. The quantitative assessment of the reproducibility using the correlation coefficients between the scan-rescan parameters are overlaid on the scatter plots.

These results show that NODDI is able to estimate the microstructure indices of neurites with a high reproducibility, comparable to DTI's ability to reproduce its less specific indices. The comparison with the results from the 4-shell protocol demonstrates additionally that the much more economical NODDI protocol does not result in a loss of the scan-rescan reproducibility, while allowing the acquisition time to be reduced to less than 30 mins.

4 Discussion

We set out to evaluate the reproducibility of NODDI, which is evident from the results presented in this work. We demonstrate that NODDI outputs precise estimates of brain microstructure, utilising the optimised protocol. The accuracy of NODDI parameters, using these protocols, has been established in [15]. Thus NODDI has the potential to help advance the research for understanding brain microstructure and developing normal brain atlases, as well as ones for specific pathologies.

However, just like DTI, NODDI parameters are also affected by drift in the scanner. A slight upward bias in the estimation of ν_{ic} can be seen in the scatter plots, by the shift of the data points above the line of equality. This can be attributed to the drift in intensities between the two scans and corresponds to the visible downward bias in the MD scatter plots (lower MD values are associated with IC regions, resulting in a higher ν_{ic} in those regions).

The correlation between the two estimates of ODI in GM is particularly low, especially compared to that in WM. This is because the dendrites have a high orientation dispersion, which leads to high ODI values that are more difficult to estimate precisely, as noted in the original work [15].

The acquisition time for NODDI protocol can potentially be reduced to as little as 10 mins by utilising fewer gradient directions in the shells, with no significant loss in accuracy of the estimates (see [15]). The reproducibility of parameters estimates using these NODDI *sub-shells* will be part of a more comprehensive reproducibility study, in future.

The results obtained show tremendous promise for utilisation of NODDI as a technique to accurately and precisely estimate microstructure. However further work is required to establish it as a clinical diagnostic tool. This would include a region based quantification of the error associated with the estimated parameters (similar to [5, 7]), which will determine the ability of the method to distinguish a patient image from that of a normal subject.

4.1 Future work

Future extensions to the work will include a more comprehensive evaluation of reproducibility of the NODDI technique and the optimised protocol, using more subjects. To evaluate the model's use as diagnostic tool for certain brain disorders, it will be important to compare the precision of NODDI parameters, with the expected differences between normal and patient brain.

The NODDI model currently captures the neurite density by a symmetric Watson model. A more comprehensive Bingham model will be incorporated in NODDI which accounts for more realistic dispersion in orientation of neurites. The NODDI model will also be implemented and available for open access within the Camino framework.

4.2 Acknowledgements

This work is supported by the future and emerging technologies (FET) program of the EU FP7 framework through the CONNECT consortium, and the MS society of Great Britain and Northern Ireland, the ISRT and the CBRC. DCA is additionally funded by ESPRC under grant EP/E007748.

References

[1] D. C. Alexander. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. *Magnetic Resonance in Medicine*, 60:439–448, 2008.

[2] D. C. Alexander, P. L. Hubbard, M. G. Hall, E. A. Moore, M. Ptito, G. J. M. Parker, and T. B. Dyrby. Orientationally invariant indices of axon diameter and density from diffusion MRI. *NeuroImage*, 52:1374–1389, 2010.

[3] Y. Assaf and Y Cohen. Inferring microstructural information of white matter form diffuion MRI. In H. Johansen-Berg and T. E. J. Behrens, editors, *Diffusion MRI: from quantitative measurement to in vivo neuroanatomy*, pages 127–146. Academic Press, 2009.

[4] P. J. Basser, J. Matiello, and D. Le Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, 66:259–267, 1994.

[5] D. Bonekamp, L. M. Nagae, M. Degaonkar, M. Matson, W. M. A. Abdalla, P. B. Barker, S. Mori, and A. HorskĞ. Diffusion tensor imaging in children and adolescents: Reproducibility, hemispheric, and age-related differences. *NeuroImage*, 34(2):733–742, 2007.

[6] J. L. Conel. *The postnatal development of the human cerebral cortex*. Harvard University Press, Cambridge, USA, 1939.

[7] L. E. Danielian, N. K. Iwata, D. M. Thomasson, and M. K. Floeter. Reliability of fiber tracking measurements in diffusion tensor imaging for longitudinal study. *NeuroImage*, 49(2):1572–80, 2010.

[8] J.C. Fiala, J. Spacek, and K.M. Harris. Review: Dendritic spine pathology: Cause or consequence of neurological disorders. *Brain Research Reviews*, 39:29–54, 2002.

[9] B. Jacobs, L. Driscoll, and M. Schall. Life-span dendritic and spine changes in areas 10 and 18 of human cortex: A quantitative golgi study. *Journal of comparative neurology*, 386:661–680, 1997.

[10] B. Jacobs, M. Schall, M. Prather, E. Kapler, L. Driscoll, S. Baca, J. Jacobs, K. Ford, M. Wainwright, and M. Treml. Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study. *Cerebral Cortex*, 11:558–571, 2001.

[11] K.V Mardia and P.E Jupp. Directional statistics. *Wiley series in probability and statistics*, 1990.

[12] C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. Di Chiro. Diffusion tensor MR imaging of the human brain. *Radiology*, 201:637–648, 1996.

[13] Diffusion Tensor Imaging ToolKit. January 2012. URL http://dti-tk.sourceforge.net.

[14] C. F Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R Kikinis. Processing and visualization for diffusion tensor MRI. *Medical Image Analysis*, 6(2):93–108, 2002.

[15] H. Zhang, T. Schneider, C.A. Wheeler-Kingshott, and D.C. Alexander. NODDI: Practical *in vivo* neurite orientation dispersion and density imaging of the human brain. *NeuroImage*, 61:1000–1016, 2012.

Interactive Segmentation of Medical Images: A Survey

Feng Zhao
Xianghua Xie
{f.zhao,x.xie}@swansea.ac.uk

Department of Computer Science
Swansea University, Singleton Park
Swansea SA2 8PP, UK

Abstract

Image segmentation is a fundamental process in most systems that support medical diagnosis, surgical planning and treatments. Generally this process is done manually by clinicians, which may be time-consuming and tedious. To alleviate the problem, a number of interactive segmentation methods have been proposed in the literature. These techniques take advantage of automatic segmentation and allow users to intervene the segmentation process by incorporating prior-knowledge, validating results and correcting errors, thus potentially lead to accurate segmentation results. In this paper, we present a survey of interactive segmentation techniques popular for medical image analysis.

1 Introduction

Due to the restrictions imposed by image acquisition, pathology, and biological variation, the medical images captured by various imaging modalities such as X-ray computed tomography (CT) and magnetic resonance imaging (MRI) are generally of high complexity and ambiguity. Image segmentation is typically used to locate objects of interest and their boundaries to make the representation of a volumetric image stack more meaningful and easier for analysis. Traditionally, this process is manually done slice by slice, which requires expert knowledge to obtain accurate boundary information for the regions of interest. This editing process may take a lot of time as well. A number of computer-aided segmentation techniques have been developed for medical images, which can usually be distinguished as automatic methods and interactive methods.

Automatic segmentation techniques such as thresholding [64], watershed [23], edge detection [48], morphological operation [32], shape analysis [19], and supervised learning [57] are usually applicable for the segmentation of well-circumscribed objects. When applied to a stack of medical images, they are able to generate rough segmentation results. These results can be further refined by the intervention of human experts. In computer-aided diagnosis, therapy planning and treatment, interactive segmentation [6, 29, 76] has become more and more popular in recent years, as the combination of human experts and machine intelligence can provide improved segmentation accuracy and efficiency with minimal user intervention [35]. The improved segmentation results can be used to reconstruct the 3D structures of tissues and enhance the real-time visualization on the screen for clinicians to navigate within the tissues freely. This can provide great benefits to many applications including locating tumors, measuring tissue volumes, surgery, and diagnosing diseases.

In this survey, we will focus on the interactive segmentation methods for medical images. Our goal is to better understand the implications of user interaction for the design of interactive segmentation methods and how they affect the segmentation results.

2 Interactive Segmentation Methodologies

Interactive segmentation [50, 63] plays an important role in the segmentation of medical images, where user intervention is suggested as an additional source of information. They leverage the expert knowledge of users to produce accurate segmentation of anatomical structures, which facilitates measurement and diagnosis of various diseases. Many approaches have been taken in interactive segmentation, which can be broadly classified into the following categories.

2.1 Fundamental approaches

In this section, we will review some common techniques (e.g., level set, region growing) that are used in interactive segmentation of medical data.

Edge-based and region-based level set segmentation methods provide a direct way to estimate the geometric properties of anatomical structures. They are popular as a general framework for many applications of medical image analysis [3, 17], such as brain MR images and 3D CT of carotid arteries. Region growing [1] is a simple region-based interactive segmentation method. Several variants of this technique have been proposed for medical image segmentation, e.g., the adaptive region growing algorithm introduced in [69]. They perform well with respect to noise and usually produce good segmentation results. However, these techniques may result in holes or over-segmentation due to noise or variation of intensity.

Statistical approach [25] is also applied to identify different tissue structures from medical images, which involves manual interaction to segment images in order to obtain a sufficiently large set of training samples. This technique is mainly applicable for problems with sufficient prior knowledge about the shape or appearance variations of the relevant structures [15, 30]. Mortensen and Barrett [49] developed an effective graphical tool (Intelligent Scissors) for performing 2D segmentation by providing immediate feedback for boundary selection as the mouse moves, which gives the user constant awareness of what belongs to the current selection. Other graph-based segmentation tools include region-based Intelligent Paint [55] and 3D Live Surface [2].

2.2 Learning-based approaches

This interactive strategy can react dynamically to the user based on the input priors (e.g., shape and appearance), and then predict the segmentation results for the user. In this framework, the user only needs to label the foreground and background on a single volumetric data, the algorithm learns the correlation between them adaptively, and completes the segmentation on other volumetric data automatically. The goal is to improve the performance of the computational part and possibly reduce the need for future user intervention, leading to interaction efficiency.

In the method described by Elliot et al. [20], the segmentation result obtained with user interaction is compared to the result obtained when the default parameter settings are used. The difference between the two is used to calibrate the parameters for the computational part, which are used as default values in future segmentation sessions. In slice-by-slice segmentation of 3D images, the information obtained with interaction in one slice can be propagated to the next in different ways. In [61], all the pixels inside the resulting object are propagated as seeds for region growing in the next slice. In the active paintbrush [44], selected points inside and outside the resulting object are propagated as 'hint' that indicate regions in the next slice where the object should (or should not) be located. The interactive method described in [11] uses a set of reference contours drawn by the user to find the

optimal parameters for an elastic-contour model using a genetic algorithm. The optimized parameters are used in all the other slices in the same or another dataset. In Yu's method [77], the resulting boundary itself is propagated as the initial contour for deformation in the next slice. In the method by Wink et al. [68], the contour in the next slice is estimated on the basis of local similarity measures of the image intensity pattern at the resulting boundary.

To overcome the application dependency, Bhanu and Fonder [5] proposed a learning-based interactive segmentation approach, in which the user can select sets of examples and counter-examples to interactively train the segmentation. The image segmentation is guided by a genetic algorithm that learns the appropriate subset and spatial combination of a collection of discriminating functions, associated with image features. The genetic algorithm encodes the discriminating functions into a functional template representation, which can be applied to the input image to produce a segmentation result. In [66], Veeraraghavan and Miller combined SVM-based active learning with GrowCut interactive segmentation to achieve a robust segmentation despite user variability with a comparable accuracy to a fully user guided segmentation with half number of user interactions on average.

The lack of labeled multimodal medical image data is a major obstacle for devising learning-based interactive segmentation tools. Transductive learning (TL) or semi-supervised learning offers a workaround by leveraging unlabeled and labeled data to infer labels for the test set given a small portion of label information. Lee et al. [36] proposed a novel algorithm for interactive segmentation using TL and inference in conditional mixture naïve Bayes models (T-CMNB) with spatial regularization constraints.

2.3 Energy minimization-based approaches

This class of segmentation methods partitions an image into different regions based on energy minimization. Among many other approaches, graph cut-based methods and deformable model-based methods are particularly popular in medical image segmentation. These techniques aim to find a global optimal solution for the boundary and region segmentation of objects in images and their performance can be efficiently improved by involving users in the process, putting users in the loop, but minimizing user input.

2.3.1 Graph cut-based approaches

Based on combinatorial optimization, graph cut [6, 59] solves the segmentation by minimizing an energy function defined on a combination of both region and boundary terms. In this approach, a graph is composed of vertices representing image pixels or voxels, and edges connecting the vertices. The graph edges are assigned some nonnegative weights or costs, and a cut is a subset of edges that partition the vertices into disjoint sets. The cost function consists of both regional and boundary information, which needs to be well defined to provide a globally optimal solution. Many current techniques use graph cut for image segmentation. It has been shown to be effective in the segmentation of images [40, 56] and volumes [2]. The use of graph cut for segmentation of 3D surfaces has been extensively validated for medical image volumes [39]. However, the execution time can be tens of minutes to cut volumes of 2-8M voxels. To accelerate the process, a single layer of oversegmentation regions has been used in the place of voxels for medical volumes which reduces the computation time to tens of seconds [78]. Lombaert et al. [43] used a resolution pyramid to perform coarse-to-fine refinement, enabling computation on the order of tens of seconds as well. In these techniques, the users are involved in the process by roughly marking out the objects of interest and the background before applying the graph cut-based segmentation. By instant feedback, additional user interaction is specified to refine the results.

2.3.2 Deformable model-based approaches

Based on variational framework, deformable modeling [13, 21, 29, 45] segments images by minimizing an energy function defined on a continuous contour or surface. It can adapt to complex shape variations and incorporate priors to regularize segmentation. Deformable modeling has been widely applied in applications such as shape extraction and object tracking, in which curves or surfaces evolve under the influence of both internal and external forces to extract the object boundaries.

Explicit models such as active contour model (Snakes) [21, 29] represent contours or surfaces in their parametric form during deformation, which have the ability to track the points on the curves or surfaces across time, and are suitable for real-time applications. However, they generally have difficulties in handling topological changes due to the parameterization of the curves or surfaces. To address these limitations, McInerney and Terzopoulos [47] developed topology adaptive deformable models by formulating deformable surfaces in terms of an affine cell image decomposition to deal with topological changes usually existing in medical image volumes. This explicit model requires a periodic reparameterization mechanism to manage complex shapes and changes in topology. This technique can effectively segment complex anatomic structures from medical volume images. However, it only performs well when the model is required to inflate or deflate everywhere, which limits its applications. New approaches [7, 18, 33] have been proposed to handle topological changes. These techniques generally involve a set of heuristic algorithms to detect self-intersections and handle splitting and merging of the deforming grid, which can be computationally expensive. In addition, they may not work well on structures consisting of complex topologies.

To address the limitations of explicit deformable models, implicit deformable models [13, 45] are introduced, based on the theory of curve evolution and the level set method [51, 58]. In the implicit models, the evolution of curves or surfaces is implicitly represented as a level set of a higher dimensional scalar function and the deformation of the models is based on geometric measures such as the unit normal and curvature. Thus, the evolution is independent of the parameterization and topological changes such as splitting and merging can be handled automatically. Implicit deformable models have been widely used in the segmentation of anatomical structures from 3D medical images [3, 28, 34].

Deformable models often vary in the object boundary representation and external force field used. Previous approaches can be distinguished as gradient-based methods [37, 45, 54, 72, 74], region-based methods [14, 17, 30, 53, 67], and hybrid methods [31, 70]. Gradient-based techniques have been found useful when there is limited prior knowledge and image gradients are reasonable indications of object boundaries. However, they require careful initialization and it may be difficult for them to achieve initialization invariance and robust convergence. This is especially true when segmenting objects with complex geometries and shapes in 3D images. Region-based techniques have been widely applied to image segmentation as well. In the popular approach [14], Chan and Vese assumed the image consists of regions of approximately piecewise-constant intensities, and then extracted the objects based on the average intensities inside and outside the contour. This method is useful for the extraction of objects with smoothly varying boundaries. However, it has difficulties dealing with image regions with intensity inhomogeneity. Other region-based approaches also assumed that the image objects are composed of distinct regional features. This is usually not true for real images due to intensity inhomogeneity and multimodal nature. In the hybrid approach [31], Kimmel used image gradient vector directions as an alignment measure, combined with the geodesic active contour and minimal variance criterion [14].

The alignment measure is used to optimize the orientation of the curve with respect to the image gradients. This measure, together with the gradient-based geodesic measure and the region-based minimal variance criterion is then used to push or pull the contour towards the image boundary. However, this hybrid technique requires careful tuning of the different parameters associated with various measures in order to efficiently bridge the image gradient and regional information. In addition, only local edge information is used in the alignment measure, while edge information of pixels located away from the contour is not considered.

The geometric active contour models [13, 45] and subsequent geodesic active contour models [12, 60] have difficulties in handling the boundary concavities, weak edges and image noise. The generalized gradient vector flow [73, 74] achieves some improvements but has convergence issues caused by saddle or stationary points in its force field. In [71, 72], Xie and Mirmehdi presented a novel edge-based model where the introduced external force field is based on the hypothesized magnetic force between the active contour and object boundaries. This method shows significant improvements in handling weak edges, broken boundaries, and complex geometries. However, its analogy based on magnetostatics cannot be directly applied to 3D or higher-dimensional images. Recently, Yeo et al. proposed a novel 3D deformable model [75, 76] based on a geometrically induced external force field, which is called the geometric potential force (GPF) field as it is based on the hypothesized interactions between the relative geometries of the deformable model and the object boundary characterized by image gradients. The evolution of the deformable model is solved using the level set method so as to facilitate topological changes automatically. The bi-directionality of the proposed GPF field allows the new deformable model to deal with arbitrary cross-boundary initializations, which is very useful in the segmentation of complex geometries, and facilitates the handling of weak image edges and broken boundaries. Moreover, the GPF deformable model can effectively overcome image noise by enhancing the geometrical interaction field with a nonlocal edge-preserving algorithm. The vector force field introduced in this work is a generalized version of the magnetic force field described in the MAC model [72], but it can be extended to higher dimensions.

3 Interactions in Medical Image Segmentation

In an interactive segmentation framework, user intervention is tightly coupled with an automatic segmentation algorithm leveraging the user's high-level anatomical knowledge and the automated method's computational capability. Real-time visualization on the screen enables the user to quickly validate and correct the automatic segmentation results in a sub-domain where the variational model's statistical assumptions do not agree with the user's expert knowledge. The user intervention mainly includes initialization of the methods, checking the accuracy of the results produced by automatic segmentation, and corrections to the segmentation results using specialized interactive segmentation tools. As shown in Table 1, interactions in the segmentation of medical images can be broadly classified into three types: pictorial input on an image grid, parameter tuning, and menu option selection. The segmentation results obtained with new configurations (e.g., mouse clicking/drawing, new parameter values, another menu option) are visualized on the screen in real time for further user evaluation.

Interactive segmentation techniques are very important for fast and reliable extraction of the regions of interest. The level of user interaction in different methods varies in terms of the amount and type of information provided by the users. Their underlying mathematical framework is a significant factor determining the form of interaction. In region growing-based methods [1, 69], the interaction is the selection of initial seed points. In the united

Interactions in Medical Image Segmentation	Examples
Pictorial input (points, lines, or regions) on an image grid	Points of background and objects [24, 26, 44] Seeds for region growing [1, 69] Point of object for initiating an inflating 3D balloon [22] Center point and radius [9] Rectangles indicating regions of interest [42] Features of different types of objects [65] Points attracting/repelling the contour [13, 21, 29] Initial curve/surface of objects of interest [46]
Parameter tuning using slider, dial, or similar interface	Scale for computing image derivatives [10, 42] Balance of weights in the cost function [8] Maximum number of iterations [10] Maximum size of segmented regions [62]
Menu option selection by mouse clicking	Accept/reject the segmentation results [65] Type of geometry model [8, 27] Properties of objects of interest [26]

Table 1: Type of interactions in the segmentation of medical images.

Snakes framework [41], the user controls the snake evolution by 'planting' seed points. The GrabCut technique [56] is based on the discrete graph-cut approach, where image pixels represent graph vertices. The partitioning of the image into object and background regions is obtained by solving the min-cut problem in graphs. The user controls the segmentation by labeling regions, which are correspondingly assigned to either the source or the sink of the graph. The selected regions provide color statistics that characterize the object and the background and are utilized for segmentation. In [52], Paragios presented a semi-automatic segmentation of the left ventricle. The method uses linear or quadratic interpolation to convert the user input into closed structures. Therefore, the feedback is not part of the level set formulation. In [38], a method applying dual-front active contours and active regions for 3D cortical segmentation is proposed. The user can modify the initialization of the active region by adding or deleting labels. A probabilistic level-set method which supports user interaction is demonstrated in [16]. The user-labeled input points are viewed as independent measurements of the scene. In [4], Ben-Zadok et al. developed a novel active-contour segmentation framework, which supports an intuitive and friendly user interaction subject to the 'bottom up' constraints introduced by the image features. Applying the level-set method [51], a fully automatic segmentation is first obtained by minimizing a cost functional that is uniquely based on the image data. The user does not 'edit' the initial segmentation, but influences its evolution with a few mouse clicks located in regions of 'disagreement'. The user input is represented as a continuous energy term that is incorporated into the primary level-set cost functional. This additional term affects the gradient descent process by attracting it toward a new local minimum, which results in a modified segmentation consistent with both the low-level image data and the top-down user feedback points.

4 Conclusion

In this paper, we briefly introduce the interactive image segmentation techniques in many medical applications. Interactive segmentation aims to achieve interaction efficiency by incorporating intelligence with automatic segmentation, leading to the ability of learning user intention and dealing with new volumetric images. To be viable for practical applications, an interactive segmentation approach should (i) minimize user interaction, (ii) minimize segmentation variability among users and (iii) be computationally fast to allow quick user editing. These concerns can be addressed by combining the machine learning techniques with interactive segmentation algorithms. Therefore, such a combined approach could provide a promising direction for accurate segmentation of medical images. A possible direction for future work could be how to efficiently learn the intention of the user so as to reduce the number of user interactions.

References

[1] R. Adams and L. Bischof. Seeded region growing. *IEEE T-PAMI*, 16(6):641–647, 1994.

[2] C. J. Armstrong, B. L. Price, and W. A. Barrett. Interactive segmentation of image volumes with live surface. *Computers & Graphics*, 31(2):212–229, 2007.

[3] C. Baillard and C. Barillot. Robust 3D segmentation of anatomical structures with level sets. In *MICCAI*, pages 236–245, 2000.

[4] N. Ben-Zadok, T. Riklin-Raviv, and N. Kiryati. Interactive level set segmentation for image-guided therapy. In *Int'l Symposium Biomedical Imaging: From Nano to Macro*, pages 1079–1082, 2009.

[5] B. Bhanu and S. Fonder. Learning based interactive image segmentation. In *ICPR*, pages 299–302, 2000.

[6] Y. Boykov and M. P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *ICCV*, pages 105–112, 2001.

[7] J. Bredno, T. M. Lehmann, and K. Spitzer. A general discrete model in two, three, and four dimensions for topology-adaptive multichannel segmentation. *IEEE T-PAMI*, 25(5):550–563, May 2003.

[8] T. A. Buck, H. H. Ehricke, W. Strasser, and L. Thurfjel. 3D segmentation of medical structures by integration of ray-casting with anatomic knowledge. *Computers & Graphics*, 19(3):441–449, 1995.

[9] A. Bzostek, et al. Isolating moving anatomy in ultrasound without anatomical knowledge: Application to computer-assisted pericardial punctures. In *MICCAI*, pages 1041–1048, 1998.

[10] J. E. Cabral, K. S. White, Y. Kim, and E. L. Effmann. Interactive segmentation of brain tumors in MR images using 3D region-growing. In *SPIE Conf. Medical Imaging*, pages 171–181, 1993.

[11] S. Cagnoni, A. B. Dobrzeniecki, R. Poli, and J. C. Yanch. Genetic algorithm-based interactive segmentation of 3D medical images. *Image and Vision Computing*, 17(12):881–895, 1999.

[12] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contour. *IJCV*, 22(1):61–79, 1997.

[13] V. Caselles, et al. A geometric model for active contours. *Numer. Math.*, 66(4):1–31, 1993.

[14] T. Chan and L. Vese. Active contours without edges. *IEEE T-IP*, 10(2):266–277, Feb. 2001.

[15] A. C. S. Chung and J. A. Noble. Statistical 3D vessel segmentation using a rician distribution. In *MICCAI*, pages 82–89, 1999.

[16] D. Cremers, O. Fluck, M. Rousson, and S. Aharon. A probabilistic level set formulation for interactive organ segmentation. In *SPIE Medical Imaging*, 2007.

[17] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape. *IJCV*, 72(2):195–215, 2007.

[18] H. Delingette and J. Montagnat. Shape and topology constraints on parametric active contours. *J. CVIU*, 83 (2):140–171, 2001.

[19] S. Diciotti, et al. Automated segmentation refinement of small lung nodules in CT scans by local shape analysis. *IEEE T-BE*, 58(12):3418–3428, 2011.

[20] P. J. Elliot, J. M. Knapman, and W. Schlegel. Interactive image segmentation for radiation treatment planning. *IBM Systems Journal*, 31(4):620–634, 1992.

[21] H. Eviatar and R. L. Somorjai. A fast simple active contour algorithm for biomedical images. *PR Letters*, 17: 969–974, 1996.

[22] J. D. Gill, et al. Development and evaluation of a semi-automatic 3D segmentation of the carotid arteries from 3D ultrasound images. In *SPIE Conf. Medical Imaging*, pages 214–221, 1999.

[23] V. Grau, et al. Improved watershed transform for medical image segmentation using prior information. *IEEE T-MI*, 23(4):447–458, April 2004.

[24] L. D. Griffin, et al. Hierarchical segmentation satisfying constraints. In *BMVC*, pages 135–144, 1994.

[25] J. Hao and M. Li. A supervised bayesian method for cerebrovascular segmentation. *WSEAS Trans. Signal Process.*, 3(12):487–495, 2007.

[26] W. E. Higgins, et al. Semi-automatic construction of 3D medical image-segmentation processes. In *SPIE Conf. Visualization in Biomedical Computing*, pages 59–71, 1994.

[27] K. P. Hinshaw, R. B. Altman, and J. F. Brinkley. Shape-based models for interactive segmentation of medical images. In *SPIE Conf. Medical Imaging*, pages 771–780, 1995.

[28] M. Holtzman-Gazit, R. Kimmel, N. Peled, and D. Goldsher. Segmentation of thin structures in volumetric medical images. *IEEE T-IP*, 15(2):354–363, Feb. 2006.

[29] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *IJCV*, 1(4):321–331, 1988.

[30] J. Kim, J. W. Fisher, A. Yezzi, M. Cetin, and A. S. Willsky. A nonparametric statistical method for image segmentation using information theory and curve evolution. *IEEE T-IP*, 14(10):1486–1502, Oct. 2005.

[31] R. Kimmel. *Geometric Level Set Methods in Imaging, Vision, and Graphics*, chapter Fast edge integration, pages 59–77. Berlin, Germany: Springer-Verlag, 2003.

[32] T. Kubota, et al. Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models. *Medical Image Analysis*, 15(1):133–154, 2011.

[33] J.-O. Lauchaud and B. Taton. Deformable model with a complexity independent from image resolution. *J. CVIU*, 99(3):453–475, 2005.

[34] M. Law and A. Chung. A deformable surface model for vascular segmentation. In *MICCAI*, pages 59–67, 2009.

[35] N. Lee, R. T. Smith, and A. F. Laine. Interactive segmentation for geographic atrophy in retinal fundus images. In *42nd Asilomar Conf. Signals, Systems and Computers*, pages 655–658, 2008.

[36] N. Lee, et al. Interactive segmentation in multimodal medical imagery using a bayesian transductive learning approach. In *Medical Imaging: Computer-Aided Diagnosis*, pages 72601W–1–10, 2009.

[37] C. Li, J. Liu, and M. Fox. Segmentation of edge preserving gradient vector flow: An approach toward automatically initializing and splitting of snakes. In *CVPR*, pages 162–167, 2005.

[38] H. Li, A. Yezzi, and L. D. Cohen. 3D brain segmentation using dual-front active contours with optional user

interaction. *Int'l J. Biomedical Imaging*, 2006:1–17, 2006.
[39] K. Li, X. Wu, D. Z. Chen, and M. Sonka. Optimal surface segmentation in volumetric images-A graph-theoretic approach. *IEEE T-PAMI*, 28(1):119–134, 2006.
[40] Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum. Lazy snapping. In *ACM SIGGRAPH*, pages 303–308, 2004.
[41] J. Liang, T. McInerney, and D. Terzopoulos. United snakes. *Medical Image Analysis*, 10(2):215–233, 2006.
[42] L. M. Lifshitz and S. M. Pizer. Multiresolution hierarchical approach to image segmentation based on intensity extrema. *IEEE T-PAMI*, 12(6):529–540, 1990.
[43] H. Lombaert, Y. Sun, L. Grady, and C. Xu. A multilevel banded graph cuts method for fast image segmentation. In *ICCV*, pages 259–265, 2005.
[44] F. Maes. *Segmentation and registration of multimodal images: From theory, implementation and validation to a useful tool in clinical practice*. PhD thesis, Katholieke Universiteit Leuven, Leuven, BE., 1998.
[45] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modelling with front propagation: A level set approach. *IEEE T-PAMI*, 17(2):158–175, Feb. 1995.
[46] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. *Medical Image Analysis*, 1(2):91–108, 1996.
[47] T. McInerney and D. Terzopoulos. Topology adaptive deformable surfaces for medical image volume segmentation. *IEEE T-MI*, 18(10):840–850, Oct. 1999.
[48] T. Mondal, A. Jain, and H. K. Sardana. Automatic craniofacial structure detection on cephalometric images. *IEEE T-IP*, 20(9):2606–2614, 2011.
[49] E. N. Mortensen and W. A. Barrett. Interactive segmentation with intelligent scissors. *Graphical Models and Image Processing*, 60:349–384, 1998.
[50] S. D. Olabarriaga and A. W. M Smeulders. Setting the mind for intelligent interactive segmentation: Overview, requirements, and framework. In *ICIPMI*, pages 417–422, 1997.
[51] S. J. Osher and J. A. Sethian. Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comp. Phys.*, 79:12–49, 1988.
[52] N. Paragios. Variational methods and partial differential equations in cardiac image analysis. In *Int'l Symposium Biomedical Imaging: From Nano to Macro*, pages 17–20, 2004.
[53] N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *IJCV*, 46(3):223–247, 2002.
[54] N. Paragios, et al. Gradient vector flow geometric active contours. *IEEE T-PAMI*, 26(3):402–407, 2004.
[55] L. J. Reese and W. A. Barrett. Image editing with intelligent paint. In *Eurographics*, pages 714–724, 2002.
[56] C. Rother, V. Kolmogorov, and A. Blake. Grabcut - Interactive foreground extraction using iterated graph cuts. In *ACM SIGGRAPH*, pages 309–314, 2004.
[57] M. Schaap, et al. Robust shape regression for supervised vessel segmentation and its application to coronary segmentation in CTA. *IEEE T-MI*, 30(11):1974–1986, 2011.
[58] J. A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Material Science*. Cambridge University Press, 1999.
[59] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE T-PAMI*, 22(8):888–905, Aug. 2000.
[60] K. Siddiqi, et al. Area and length minimizing flows for shape segmentation. *IEEE T-IP*, 7(3):433–443, 1998.
[61] J. Sijbers, et al. Volume quantization of the mouse cerebellum by semi-automatic 3D segmentation of magnetic resonance images. In *SPIE Conf. Medical Imaging*, pages 553–560, 1996.
[62] G. J. Sivewright and P. J. Elliot. Interactive region and volume growing for segmenting volumes in MR and CT images. *Medical informatics*, 19(1):71–80, 1994.
[63] A. Smeulders, et al. Design considerations for interactive segmentation. In *CVIS*, pages 5–12, 1997.
[64] C. M. Smith, et al. Automatic thresholding of three-dimensional microvascular structures from confocal microscopy images. *J. Microscopy*, 225(3):244–257, 2007.
[65] J. K. Udupa, et al. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE T-MI*, 16(5):598–609, 1997.
[66] H. Veeraraghavan and J. V. Miller. Active learning guided interactions for consistent image segmentation with reduced user interactions. In *Int'l Sym. Biomedical Imaging: From Nano to Macro*, pages 1645–1648, 2011.
[67] L. Vese and T. Chan. A multiphase level set framework for image segmentation using the mumford and shah model. *IJCV*, 50(3):271–293, 2002.
[68] O. Wink, K. J. Zuiderveld, and M. A. Viergever. Interactive volume segmentation using local similarity measurements. In *Computer Assisted Radiology Conf.*, page 996, 1997.
[69] J. Wu, F. Ye, J. Ma, X. Sun, J. Xu, and Z. Cui. The segmentation and visualization of human organs based on adaptive region growing method. In *Int'l Conf. Comp. Inf. Tech.*, pages 439–443, 2008.
[70] X. Xie and M. Mirmehdi. Rags: Region-aided geometric snake. *IEEE T-IP*, 13(5):640–652, May 2004.
[71] X. Xie and M. Mirmehdi. Magnetostatic field for the active contour model: A study in convergence. In *BMVC*, pages 127–136, 2006.
[72] X. Xie and M. Mirmehdi. Mac: Magnetostatic active contour model. *IEEE T-PAMI*, 30(4):632–647, 2008.
[73] C. Xu and J. L. Prince. Generalized gradient vector flow external forces for active contours. *Signal Processing*, 71(2):131–139, 1998.
[74] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE T-IP*, 7(3):359–369, Mar. 1998.
[75] S. Y. Yeo, X. Xie, I. Sazonov, and P. Nithiarasu. Geometric potential force for the deformable model. In *BMVC*, pages 1–11, 2009.
[76] S. Y. Yeo, X. Xie, I. Sazonov, and P. Nithiarasu. Geometrically induced force interaction for three-dimensional deformable models. *IEEE T-IP*, 20(5):1373–1387, May 2011.
[77] X. Yu and J. Yla-Jaaski. Interactive surface segmentation for medical images. In *Int'l Conf. Signal Processing*, pages 1178–1181, 1996.
[78] X. Yuan, et al. Volume cutout. *Visual Computer (Special Issue of Pacific Graphics)*, 21(8-10):745–754, 2005.

Advanced Blur Removal Methods with Applications to Retinal Imaging for Ophthalmology

Bryan Williams[1]
bryan@liv.ac.uk
Ke Chen[1]
www.liv.ac.uk/cmit
Yalin Zheng[2]
cmit@liv.ac.uk
Simon Harding[2]
cmit@liv.ac.uk

[1] Department of Mathematical Sciences
The University of Liverpool
Liverpool, UK

[2] Department of Eye and Vision Science
The University of Liverpool
Liverpool, UK

Abstract

Blurring of images occurs in many fields, causing significant problems in retinal imaging. In any diabetic retinopathy screening programme, up to 10% of the images are ungradable due to inadequate clarity or poor field definition. It is important to obtain as much information as possible of retinal vessels and other structures.

Deblurring is a major technique that may be developed to restore the lost true image. We present non-blind and blind approaches to tackle this problem, including an approach for multi-channel images. We propose a new solution algorithm for removing unknown blur (blind case) and show results for retinal images.

1 Introduction

Image processing techniques, such as image reconstruction which includes removing image noise from a given image (denoising) [12], reconstructing an image from an given blurred image (deblurring) [6], reconstructing the missing or damaged portion of an image (inpainting) [3], emphasizing the boundaries of an image by different filters or segmenting a image into subregions (segmentation) [5], have been widely used in many areas. Despite significant development in photographic techniques and technology, blur is still a major cause for image quality degradation in clinical settings. While new segmentation models can cope with noise, they become inflective for blurred images. This is due to many factors such as motion of the camera or more commonly in the case of retinal images the target scene, defocusing of the lens system, imperfections in the electronic, photographic, transmission medium, or obstructions.

An observed blurred image can be written as a convolution of the true image with a Linear Shift-Invariant (LSI) blur, known as the Point Spread Function (PSF) or spatial invariant/variant unknown kernel K [9].

There are three main deconvolution problems: (1) blind deconvolution, which includes the cases when both kernel and image are unknown, (2) semi-blind restoration,

in which the kernel is assumed to belong to a class of parametric functions, or (3) non-blind deconvolution where only the image is unknown. All three types are important not only in many scientific applications such as astronomical imaging, medical imaging, and remote sensing, but also for consumer photography.

Deconvolution in the case of known blur, assuming the linear degradation model, has been investigated widely in the last few decades giving rise to a variety of solutions [2, 10, 13]. In non-blind deconvolution, the point spread function is assumed known even though this information is not available in most of the real applications. In cases when the blur is not known the problem becomes harder and much more challenging. Blind deconvolution, which is our main concern below, was first introduced by [6] and a lot of work has been carried out so far to improve the model [1, 7, 11] and an excellent tutorial has been provided by [9].

This paper is organised as follows: in Section 2, we present the formulation for non-blind deblurring and the splitting idea with some results. In Section 3 we show the formulation of the blind method with results, including applications in retinal imaging. In the Section 4, we give the conclusion.

2 Non-Blind Deblurring

The idea of minimising an energy functional of the form

$$\min_u \left\{ ||k*u-z||_2^2 + \alpha \int_\Omega L(u) d\Omega \right\} \tag{1}$$

where the first term is the least squares term which aims to keep the restored image as close to the true image as possible and the second term is a regularisation term which aims to restore edges lost in the reconstruction, was proposed by [12]. A number of functions are commonly selected for $L(u)$, such as $L(u) = u^2$ (Tikhonov [14]) or $L(u) = |\nabla u|^2$. The Total Variation (TV) regularisation term given by $L(u) = |\nabla u|$ has been widely used due to its effectiveness with preserving edges. Minimising equation (1) we derive the Euler Lagrange equation

$$K^T Ku - K^T z - \alpha \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) = 0,$$

where K is an $n^2 \times n^2$ dense matrix of structured blocks. To cope with $\nabla u = 0$, we modify $|\nabla u|$ by $|\nabla u|_\beta = \sqrt{(\nabla_u)^2 + \beta}$ for $\beta > 0$. We therefore aim to solve

$$K^T Ku - K^T z - \alpha \nabla \cdot \left(\frac{\nabla u}{|\nabla u|_\beta} \right) = 0. \tag{2}$$

To solve this non-linear partial differential equation (PDE) we use iterative methods. Time marching is effective but typically proves to be slow to obtain good results. Instead, we use a Conjugate Gradient method aided by preconditioners which aim to increase the stability and speed of the system, such as the Product Preconditioner [15] or the Cosine Transform Preconditioner [4].

2.1 Splitting Deblurring and Denoising

We modify the functional by replacing the restored image variable in the least squares term with a new variable, distinct from that used in the regularisation term, and add a further term to minimise the difference between them. Our modified functional is given by

$$f(u,v) = \frac{1}{2}||k*v-z||_2^2 + \alpha \int_\Omega |\nabla u| d\Omega + \frac{\gamma}{2}||u-v||_2^2. \tag{3}$$

Minimisation of equation (3) with respect to u and v yields the Euler Lagrange equations

$$K^T Kv - K^T z - \gamma(u - v) = 0 \tag{4}$$

$$\gamma(u - v) - \alpha\nabla \cdot \left(\frac{\nabla u}{|\nabla \tilde{u}|_\beta} \right) = 0 \tag{5}$$

which we solve for v and u respectively [8]. We may use fixed point and the 2-dimensional fast Fourier transform (fft) to solve equation (4) for v. In order to solve equation (5) for u we use Time Marching or Conjugate Gradient. We make an initial estimate u of the true image u_{true}, which we typically take to be the received image z and repeatedly solve equations (4) and (5) for u and v respectively until u and v are sufficiently close. Our algorithm is given as

Algorithm 1 Non-blind Deblurring

procedure NONBLINDSPLITTING($z, k, \alpha, \gamma, tol$)
  $u \leftarrow z$
  repeat
    Solve $(K^T K + \gamma I)v = K^T z + \gamma u$ for v
    Solve $\gamma(u - v) - \alpha\nabla \cdot \left(\frac{\nabla u}{|\nabla \tilde{u}|_\beta} \right) = 0$ for u
  until $||u - v||_2^2 \leq tol$
end procedure

2.2 Experimental Results

We present experimental results of this method in the non-blind case using the satellite and retinal image examples corrupted by motion and Gaussian blur in figures 1—3.

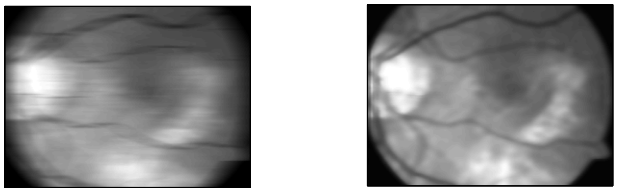


Figure 1: Retina image with motion blur. The PSNR increases from 22.847 in the received image to 27.145 in the restored image. The CPU time to obtain the restored image is 1.48.

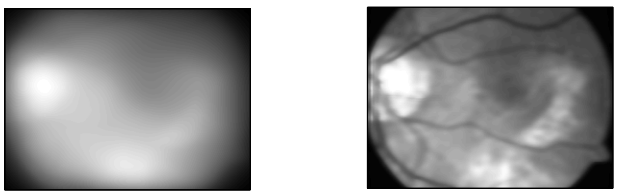


Figure 2: Retina image with strong Gaussian blur. The PSNR increases from 17.862 in the received image to 27.003 in the restored image. The CPU time to obtain the restored image is 1.63.

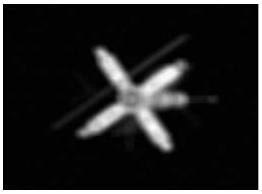
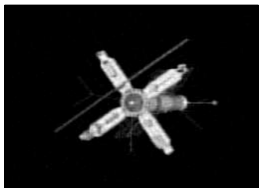



Figure 3: Satellite image with motion blur. The PSNR increases from 142.688 in the received image to 147.510 in the restored image.

3 Blind Deblurring

A model for Blind Deblurring, given by minimising the functional

$$f(u,k) = \frac{1}{2}||k*u-z||_2^2 + \alpha_1 \int_\Omega |\nabla u| d\Omega + \alpha_2 \int_\Omega |\nabla k| d\Omega. \tag{6}$$

with respect to u and k and solving via an alternate minimisation scheme was proposed by [6]. The model was shown to give excellent results for model problems, but not for testing general images where convergence is an issue. We modify this functional, splitting the restored image and kernel terms as follows:

$$f(u,v,k,h) = \frac{1}{2}||h*v-z||_2^2 + \alpha_1 \int_{\Omega_1} |\nabla u| d\Omega_1 + \alpha_2 \int_{\Omega_2} |\nabla k| d\Omega_2 + \frac{\gamma_1}{2}||u-v||_2^2 + \frac{\gamma_2}{2}||k-h||_2^2.$$

Minimising with respect to u, v, k and h respectively, we obtain the following Euler Lagrange equations

$$\gamma_1(v(x,y)-u(x,y)) + h(-x,-y)*(h(x,y)*v(x,y)-z(x,y)) = 0 \tag{7}$$

$$\gamma_1(u(x,y)-v(x,y)) - \alpha_1 \nabla \cdot \left(\frac{\nabla u(x,y)}{|\nabla u(x,y)|_\beta} \right) = 0 \tag{8}$$

$$\gamma_2(h(x,y)-k(x,y)) + v(-x,-y)*(v(x,y)*h(x,y)-z(x,y)) = 0 \tag{9}$$

$$\gamma_2(k(x,y)-h(x,y)) - \alpha_2 \nabla \cdot \left(\frac{\nabla k(x,y)}{|\nabla k(x,y)|_\beta} \right) = 0 \tag{10}$$

An overall algorithm is given in Algorithm 2.

3.1 Experimental Results

We present experimental results of the blind restoration using this model of the satellite and retina images corrupted by motion blur in figures (4) and (5).

4 Conclusions

A new splitting method algorithm is proposed for blind deconvolution restoration. Test results in retinal images are encouraging and show that it is potentially useful in medical imaging.

Algorithm 2 Blind Deblurring

function BLINDSPLITTING($z, k_{initial}, \alpha_1, \alpha_2, \gamma_1, \gamma_2, tol, maxit$)
  $u \leftarrow z$
  $k \leftarrow k_{initial}$
  for $i \leftarrow 1$ **to** $maxit$ **do**
    repeat
      Solve $u(-x,-y)uk - u(-x,-y)z + \gamma(k-h) = 0$ for k
      Solve $\gamma(h-k) - \alpha\nabla\cdot\left(\frac{\nabla h}{|\nabla\tilde{h}|}\right) = 0$ for h
    until $||h-k||_2^2 \leq tol$
    Impose: $h(x,y) \leftarrow 0$ if $h(x,y) < 0$
    Impose: $h(x,y) \leftarrow (h(x,y) + h(-x,-y))/2 \; \forall x,y \in \Omega$
    Impose: $h \leftarrow h / \int_\Omega h(x,y)dxdy$
    repeat
      Solve $k(-x,-y)kv - k(-x,-y)z + \gamma(v-u) = 0$ for v
      Solve $\gamma(u-v) - \alpha\nabla\cdot\left(\frac{\nabla u}{|\nabla\tilde{u}|}\right) = 0$ for u
    until $||u-v||_2^2 \leq tol$
    Impose: $u(x,y) \leftarrow 0$ if $u(x,y) < 0$
  end for
end function

References

[1] Mariana S. C. Almeida and Luís B. Almeida. Blind and semi-blind deblurring of natural images. *IEEE Transactions on Image Processing*.

[2] L. Bar, N. Sochen, and N. Kiryati. Semi-blind image restoration via Mumford-Shah regularization. *IEEE Transactions on Image Processing*, 15(2):483—493, Feb 2006.

[3] C. Brito-Loeza and K. Chen. Multigrid method for a modified curvature driven diffusion model for image inpainting. *Journal of Computational Mathematics*, 26 (6):856—875, 2008.

[4] Raymond H. Chan, Tony F. Chan, and Chiu Kwong Wong. Cosine transform based preconditioners for total variation deblurring. 1995.

[5] Tony F. Chan and Luminita A. Vese. Active contours without edges, 1998.

[6] Tony F. Chan and C. K. Wong. Total variation blind deconvolution. *IEEE Transactions on Image Processing*, 7(3):259—285, March 1998.

[7] Tony F. Chan and C. K. Wong. Convergence of the alternating minimization algorithm for blind deconvolution. *Linear Algebra Appl*, 316:259—285, 2000.

[8] Yumei Huang, Michael K. Ng, and You-Wei Wen. A fast total variation minimization method for image restoration. *SIAM*, 2008.

[9] D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13(3):43—64, may 1996.

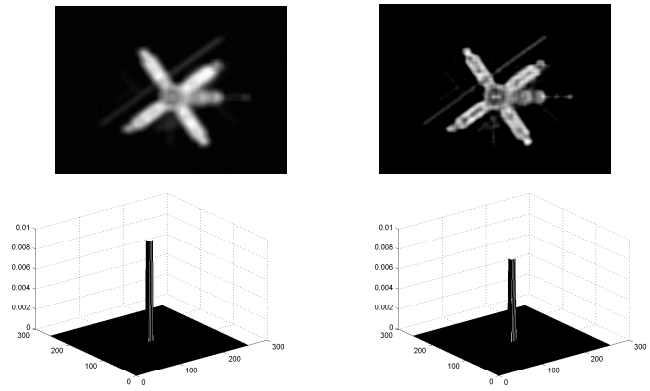


Figure 4: Satellite image corrupted with motion blur. The CPU time to obtain the restored image is 16.71. On the top left we show the received image, on the top-right is the restored image, bottom left is the initial estimate of the kernel and bottom right is the restored kernel.

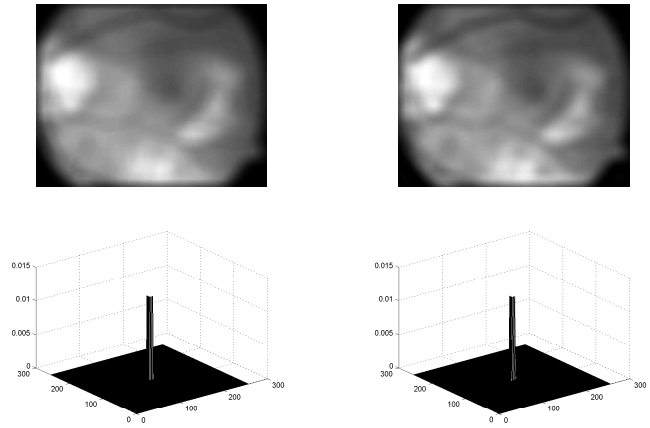


Figure 5: Retina image corrupted with motion blur. The CPU time to obtain the restored image is 18.77. On the top left we show the received image, on the top-right is the restored image, bottom left is the initial estimate of the kernel and bottom right is the restored kernel.

[10] R. L. Lagendijk, I. Biemond, and D. E. Boekee. Regularized iterative image restoration with ringing reduction. *IEEE Trans. Acoust. Speech Signal Processing*, 36(12): 1874—1888, 1988.

[11] R. Molina, J. Mateos, and A.K. Katsaggelos. Blind deconvolution using a variational approach to parameter, image, and blur estimation. *IEEE Transactions on Image Processing*, 15(12):3715—3727, Dec 2006.

[12] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. 1992.

[13] M.I. Sezan and H.J. Trussell. Prototype image constraints for set-theoretic image restoration. *IEEE Transactions on Signal Processing*, 39(10):2275—2285, oct 1991.

[14] Curtis R. Vogel. *Computational Methods for Inverse Problems*. SIAM, 2002.

[15] Curtis R. Vogel and Mary E. Oman. Fast, robust total variation-based reconstruction of noisy, blurred images. 1998.

Author Index

MUA 2012

SWANSEA

ISBN 1-901725-45-6