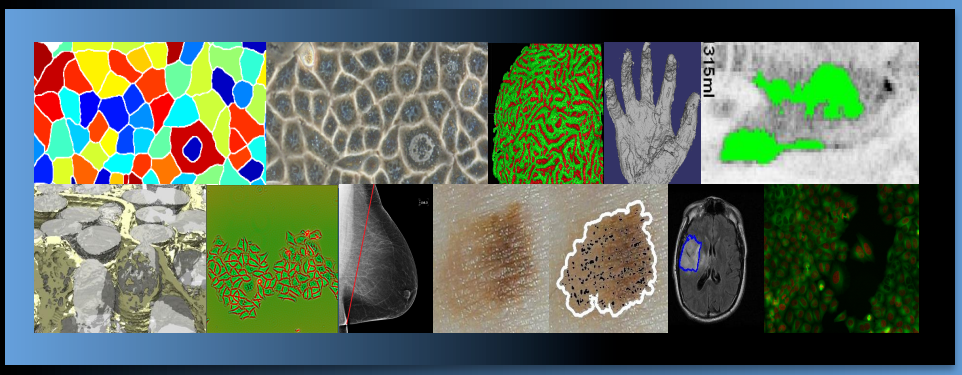
# Medical Image Understanding and Analysis 2014

Editors:
Constantino Carlos Reyes-Aldasoro
Greg Slabaugh




Proceedings of the 18th Conference on Medical Image Understanding and Analysis.

9-11 July 2014

London, United Kingdom

# Medical Image Understanding and Analysis 2014

## Sponsored by

Constantino Carlos Reyes-Aldasoro, Greg Slabaugh, Proceedings Editors
School of Mathematics, Computer Science, and Engineering
City University London
London EC1V 0HB
United Kingdom

ISBN: 1-901725-51-0

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted in any form or by means, with the prior permission in writing from the publishers, or in the reprographic reproduction in accordance with the terms of licenses issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

©BMVA July 2014

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

The publisher makes no representation, express or implied, with regard to the accuracy of the information published in this book and cannot accept any legal responsibility for any errors of omission that may be made.

Printed and bound in the United Kingdom

www.miua.org.uk

## Preface

MIUA 2014 is the eighteenth in the series of annual meetings. Since its inauguration, in 1997 at Oxford, this multidisciplinary event has been providing a forum for presenting and discussing research related to medical image analysis. The areas covered by MIUA include computer science, mathematics, engineering and physics as well as biosciences, medical research and clinical practice. The principal research interest of the conference is in methods of analysis that extract meaningful and quantitative information from images to aid diagnosis and therapy or to support research in fundamental biomedical sciences. Over the lifetime of the MIUA conferences we have seen significant advances made in the development of novel imaging modalities and methods. Many ideas first proposed by members of the medical imaging community have progressed from the research laboratory to clinical practice and are making direct impact on patient care. We are very pleased to host a conference in London that contributes to these endeavours. This years keynote lectures are delivered by three eminent academics: Professor David Hawkes from University College London, Professor Jean-Christophe Olivo-Marin from Institut Pasteur, and Professor Roger Gunn from Imperial College London. We are very grateful for their contributions.

The conference prides itself in providing a friendly forum and support for research students and young scientists. The organisers of the MIUA 2012 conference at Swansea initiated the idea of pre- conference tutorials aimed at introducing the participants to novel or emerging modalities or techniques by a leading expert in a field. This year we are grateful to Dr Jasmina Lazic of the Mathworks for a tutorial on MATLAB for Medical Image Processing, and Dr Antonio Criminisi, from Microsoft Research for a tutorial on Decision Forests in Medical Image Analysis.

MIUA was originally conceived as a UK event, however, over the years international contributions from Europe and beyond have been increasing. This year we warmly welcome participants and contributors from France, Germany, Belgium, Portugal, Italy, Czech Republic, Greece, Cyprus, Austria, USA, South Africa, Saudi Arabia, Indonesia, Taiwan, and the Republic of Korea.

There are many people whose effort and commitment contributed to the organisation of this conference and who deserve special thanks:

- The MIUA Steering Committee chaired by Bill Crum, for their unfailing support and advice.
- The reviewers, for their thoughtful comments and timely submission of paper reviews.
- Emma Leaver, Louise Gordon, Kristie Loutsiou, and Razina Patel for enthusiastic, and frequent, help with administration.
- Jack Lewis and Stephanie Schaal for maintaining the MIUA 2014 website.
- All the students, staff and session chairs, for their help before and during the conference.
- Professor Roger Crouch, for formally opening the conference.
- The British Machine Vision Association, for sponsoring bursaries for the best student papers.

- The School of Mathematics, Engineering, and Computer Science at City University London for sponsoring the conference.
- John Wiley & Sons, for sponsoring prizes.
- The British Association for Cancer Research, for sponsoring a best student prize in cancer.

Finally, many thanks go to the authors and presenters of the papers and to all the conference delegates for their scientific contributions, and through participation, for helping to maintain a healthy and vibrant medical image analysis community.

*Constantino Carlos Reyes-Aldasoro*
*Greg Slabaugh*
Co-chairs, MIUA 2014

**Chairs**

Constantino Carlos Reyes-Aldasoro, City University London
Greg Slabaugh, City University London

**Steering committee**

Ela Claridge, University of Birmingham
Bill Crum, King's College London
Tryphon Lambrou, University of Lincoln
Stephen McKenna, University of Dundee
Graeme Penney, King's College London
Nasir Rajpoot, University of Warwick
Constantino Carlos Reyes-Aldasoro, City University London
Greg Slabaugh, City University London
Xianghua Xie, Swansea University
Xujiong Ye, University of Lincoln
Reyer Zwiggelaar, Aberystwyth University

**Technical committee**

Charith Abhayaratne, University of Sheffield
Maria Algorri, Hochschule Ulm
Muhammad Asad, City University London
Li Bai, University of Nottingham
Paul Barber, Oxford University
Sarah Barman, Kingston University
Phillip Birch, University of Sussex
Abhir Bhalerao, University of Warwick
Djamal Boukerroui, Universit de Technologie de Compigne
Ela Claridge, University of Birmingham
William Crum, King's College London
Eddie Edwards, Imperial College London
Khemraj Emrith, University of the West of England
Mark Fisher, University of East Anglia
Jim Graham, The University of Manchester
Vicente Grau, University of Oxford
Benjamin Gutierrez-Becker, Technische Universitt Mnchen
Hakan Haberdar, University of Houston, Texas
Yulia Hicks, Cardiff University
Andrew King, King's College London
Fred Labrosse, University of Aberystwyth
Tryphon Lambrou, University of Lincoln
Gabriel Landini, University of Birmingham
Emma Lewis, University of Surrey
Qingde Li, University of Hull
Hongying Meng, Brunel University
Jamie McClelland, University College London
Stephen McKenna, University of Dundee
Majid Mirmehdi, University of Bristol
Emmanouil Moschidis, University of Manchester

Hammadi Nait-Charif, University of Bournemouth
Tim Nattkemper, Universitat Bielefeld
Alison Noble, University of Oxford
Boguslaw Obara, Durham University
Graeme Penney, King's College London
Michael Phillips, City University London
Andre Possani, Instituto Tecnologico Autonomo de Mexico
Kashif Rajpoot, University of Oxford
Nasir Rajpoot, University of Warwick
Constantino Carlos Reyes-Aldasoro, City University London
Victor Sanchez, Warwick University
Alberto Santamaria-Pang, G.E.
Gerald Schaefer, Loughborough University
Greg Slabaugh, City University London
Iain Styles, University of Birmingham
Joao Tavares, Universidade do Porto
Cagatay Turkay, City University London
Darren Treanor, NHS
Mohammad Uddin, Jahangirnagar University
Frank Vidal, University of Bangor
Xianghua Xie, Swansea University
Peter Weller, City University London
Guang Yang, St. George's Hospital
Xujiong Ye, University of Lincoln
Yinyin Yuan, Institute for Cancer Research
Yalin Zheng, University of Liverpool
Diwei Zhou, University of Brighton
Huiyu Zhou, Queen's University Belfast
Reyer Zwiggelaar, University of Aberystwyth

**KEYNOTE SPEAKERS**

**David Hawkes**
*Multi-Scale Image Guided Interventions*..... 2

**Jean-Christophe Olivo-Marin**
*Quantitative Cell Dynamics*..... 3

**Roger Gunn**
*Advances in the Analysis of Molecular Neuroimaging Data*..... 4

**NOVEL IMAGING METHODS AND APPLICATIONS**

**Samuele Fiorini, Lucia Ballerini, Emanuele Trucco, Alfredo Ruggeri**
*Automatic Generation of Synthetic Retinal Fundus Images*..... 7

**Julien Pansiot, Lionel Reveret, Edmond Boyer**
*Combined Visible and X-Ray 3D Imaging*..... 13

**Khemraj Emrith, Greg Slabaugh, Andy Poullis, Chris Groves, Melvyn Smith**
*Photometric Stereo Reconstruction for Surface Analysis of Mucosal Tissue*..... 19

**NEUROLOGICAL IMAGE ANALYSIS**

**Ana Ineyda Namburete, Richard Stebbing, J. Alison Noble**
*Diagnostic Plane Extraction from 3D Parametric Surface of the Fetal Cranium*..... 27

**Maria Valdes-Hernandez, Jaeil Kim, Ian Whiteford, Xinyi Qiu, Joanna Wardlaw, Jinah Park**
*Automatic Hippocampal Multimodal Assessment for Studies of Stroke and Small Vessel Disease*..... 33

**Claire Cury, Joan Glaunès, Marie Chupin, Olivier Colliot**
*Fast Template-based Shape Analysis Using Diffeomorphic Iterative Centroid*.... 39

**ANALYSIS OF CELLULAR IMAGES**

**Nicolas Jaccard, Nicolas Szita, Lewis Griffin**
*Trainable Segmentation of Phase Contrast Microscopy Images Based on Local Basic Image Features Histograms*..... 47

**Tianjun Huang, Christian Hacker, John Lucocq, Stephen McKenna**
*Detecting and Segmenting Nanodiscs in Immuno-Electron Micrographs*..... 53

**Neil Fordyce, Stephen McKenna, Christian Hacker, John Lucocq**
*Finding Golgi Stacks in Electron Micrographs*..... 59

**CANCER IMAGING**

**Jingjing Zhang, Stephen McKenna, Jianguo Zhang, Maria Coats, Frank Carey**
*Analysing the Surface Morphology of Colorectal Polyps: Differential Geometry and Pit Pattern Prediction* ..... 67

**Spyridon Bakas, Katerina Chatzimichail, Bastien Labbé, Gordon Hunter, Paul Sidhu, Dimitrios Makris**
*Fast Segmentation of Focal Liver Lesions in Contrast-Enhanced Ultrasound Data* ..... 73

**Wenda He, Minnie Kibiro, Arne Juette, Peter Hogg, Erika Denton, Reyer Zwiggelaar**
*A Novel Breast Image Preprocessing For Full Field Digital Mammographic Segmentation and Risk Classification* ..... 79

**Maria João Vasconcelos, Luis Rosado**
*Automatic Reflection Detection on Dermatological Images Acquired via Mobile Devices* ..... 85

**Damilola Okuboyejo, Oludayo Olugbara, Solomon Odunaike**
*Unsupervised Restoration of Hair-Occluded Lesion in Dermoscopic Images* ..... 91

**SEGMENTATION**

**Mohammadreza Soltaninejad, Tryphon Lambrou, Adnan Qureshi, Nigel Allinson, Xujiong Ye**
*A Hybrid Method for Haemorrhage Segmentation in Trauma Brain CT* ..... 99

**Sri Purwani, Tim Cootes, Carole Twining**
*Ensemble Registration: Combining Groupwise Registration and Segmentation* ..... 105

**Fabian Schenk, Martin Urschler, Christoph Aigner, Imme Roesner, Philipp Aichinger, Horst Bischof**
*Automatic Glottis Segmentation from Laryngeal High-Speed Videos Using 3D Active Contours* ..... 111

**Marco Körner, Mahesh Krishna, Herbert Süße, Wolfgang Ortmann, Joachim Denzler**
*Regularized Geometric Hulls for Nucleus Segmentation from Microscopy Images* ..... 117

**FEATURE ANALYSIS AND CLASSIFICATION**

**Jan Margeta, Antonio Criminisi, Daniel Lee, Nicholas Ayache**
*Recognizing Cardiac Magnetic Resonance Acquisition Planes* ..... 125

**Shan Khazendar, Jessica Farren, Hisham Al-Assam, Hongbo Du, Ahmed Sayasneh, Tom Bourne, Sabah Jassim**
*Automatic Identification of Early Miscarriage Based on Multiple Features Extracted From Ultrasound Images* ..... 131

**CHALLENGE ABSTRACT**

**Francesca Pizzorni-Ferrarese, Jonas Larsson**
*Challenges in Functional Imaging of the Visual Cortex: Data-driven Model-agnostic Estimation of Population Receptive Fields* ..... 139

**POSTER SESSION 1: SEGMENTATION AND SHAPE ANALYSIS**

**Omar Al Okashi, Hongbo Du, Joanne Selway, Chris Lelliott, Simon Maguire, Dave Melvin, Kenny Langlands, Hisham Al-Assam**
*Automatic Ribs Segmentation and Counting From Mouse X-ray Images*..... 143
**Alaa Khadidos, Victor Sanchez, Chang-Tsun Li**
*Active Contours with Weighted External Forces for Medical Image Segmentation*..... 149
**Lei Zhang, Mark Fisher, Wenjia Wang**
*Retinal Vessel Segmentation using Gabor Filter and Textons*..... 155
**Adnan Qureshi, Vitaly Schetinin**
*Computer-Aided Segmentation and Estimation of Indices in Brain CT Scans*..... 161
**Musab Sahrim, Mark Nixon, Roxana Carare**
*On Bloodvessel Branching Analysis for the Detection of Alzheimer's Disease*.... 167
**Michael Phillips, Sally Barrington, Derek Hill, Paul Marsden**
*Comparison of Threshold-Based Segmentation Methods on Pre- and Post-Therapy PET Scans*..... 173
**Carl Nelson, Martin Dixon, Philippe Laissue, Boguslaw Obara**
*Speeding Up Active Mesh Segmentation by Local Termination of Nodes*..... 179
**Jindřich Soukup, Michal Lašan, Filip Šroubek**
*Segmentation of Individual Cells in Phase Contrast Microscopy Images*..... 185
**Thomas Nketia, Jens Rittsher, J. Alison Noble**
*Utilizing Phase Retardation Features for Segmenting Cells in Phase Contrast Microscopy Images*..... 191
**Da-Chuan Cheng, Shing-Hong Liu**
*Automated Vessel Boundary Detection Using 3D Expansion of Dynamic Programming*..... 197
**Jahanzeb Ahmad, Jiuai Sun, Lyndon Smith, Melvyn Smith**
*Non-Contact Pulmonary Functional Testing Through an Improved Photometric Stereo Approach*..... 203
**Matthaios Doulgerakis, Vasileios Argyriou, George Tzimiropoulos, Barbara Pierscionek**
*Non-Intrusive Action Symmetry Measurement for Dementia and Rehabilitation*..... 209
**Radim Kolar, Petra Podlipna**
*Analysis of Images from Optical Coherence Tomography for the Nanoparticle Transport*..... 215

**POSTER SESSION 2: MACHINE LEARNING, SIMULATION, AND IMAGE ANALYSIS APPLICATIONS**

**Saad Al-Baddai, Karema Al-Subari, Ana Maria Tomé, Gregor Volberg, Elmar Lang**
*Combining EMD with ICA to Analyze Combined EEG-fMRI Data* ..... 223

**Thomas Janssens, Ine Vanhees, Jan Gunst, Helen Owen, Greet Van den Berghe, Fabian Guiza Grandas**
*Automated Histological Quantification of Trabecular Bone Tissue in Critical Illness* ..... 229

**Simon Bowyer, Chryso Kanthou, Constantino Reyes-Aldasoro**
*Analysis of Capillary-like Structures Formed by Endothelial Cells in a Novel Organotypic Assay Developed from Heart Tissue* ..... 235

**Rachel Flight, Gabriel Landini, Iain Styles, Richard Shelton, Michael Milward, Paul Cooper**
*Semi-Automated Cell Counting in Phase Contrast Images of Epithelial Monolayers* ..... 241

**Jakob Unger, Maria Schuster, Dietmar Hecker, Bernhard Schick, Joerg Lohscheller**
*A Wavelet-based Approach for Analyzing Non-Stationary Phonation Sequences from Endoscopic High-Speed Videos* ..... 247

**Akadej Udomchaiporn, Frans Coenen, Vanessa Sluming, Marta Garcia-Finana**
*3-D MRI Brain Scan Classification of Epilepsy versus Non-Epilepsy* ..... 253

**Mohammadreza Soltaninejad, Xujiong Ye, Guang Yang, Nigel Allinson, Tryphon Lambrou**
*Brain Tumour Grading in Different MRI Protocols using SVM on Statistical Features* ..... 259

**Ting Yue Yu, Hamid Deghghani, Keith Brain, Fahima Syeda, Andrew Holmes, Paulus Kirchhof, Larissa Fabritz**
*New Optical Mapping Design and Automated Algorithms for Cardiac Electrophysiology* ..... 265

**Ioannis Constantinou, Marios Pattichis, Chrysa Tziakouri, Constantinos Pattichis, Styliani Petroudi**
*Multiscale AM-FM Models and Instantaneous Amplitude Evaluation for Mammographic Density Classification* ..... 271

**Chris Meah, Richard Marshall, Ela Claridge, Kai Bongs, Iain Styles**
*A Simulation Toolbox For Biomedical Plenoptic Imaging* ..... 277

**Joaquin Correa, Manfred Auer, David Skinner**
*Integrated Tools for Next Generation Bioimaging* ..... 283

# Keynote speakers

# Professor David Hawkes

# University College London

## Multi-Scale Image Guided Interventions

Professor David Hawkes graduated from Oxford with a BA in Natural Sciences (Physics) in 1974 and obtained an MSc in Radiolobiology in Birmingham in 1975. His first post was as a clinical scientist in nuclear medicine at Southampton General Hospital. After obtaining his PhD in 1982 at the University of Surrey and the Institute of Cancer Research he spent seven years working as a clinical scientist at St. George's Hospital, London, before returning in 1988 to academia at the United Medical and Dental Schools of Guy's and St Thomas' Hospitals (UMDS), where he received a personal Chair in 1998.

Dave Hawkes is currently the Director of the Centre for Medical Image Computing, previously having been Director of the EPSRC and MRC funded Interdisciplinary Research Collaboration on Medical Images and Signals (MIAS-IRC), an £8M six year programme, from 2003 to 2007 and Chairman of the Division of Imaging Sciences at KCL (2002-2004). He spent 10 years working as a clinical scientist within the NHS before returning to academia. He is co-Founder of IXICO Ltd. (www.ixico.com), a university spin-out that provides imaging solutions to the pharmaceutical industry. His current research interests encompass image matching, data fusion, visualisation, shape representation, surface geometry and modelling tissue deformation promoting medical imaging as an accurate measurement tool and image guided interventions.

He was elected FMedSci in 2011, FBIR in 2007, FREng in 2002, FInstP in 1997 and FIPEM in 1993. He became an NIHR Senior Investigator in 2009.

# Professor Jean-Christophe Olivo-Marin

## Institut Pasteur

### Quantitative Cell Dynamics

Jean-Christophe Olivo-Marin, PhD, is the head of the BioImage Analysis Unit and the chair of the Cell Biology and Infection Department at Institut Pasteur, Paris. He was a cofounder and Chief Technology Officer of the Institut Pasteur Korea, Seoul. His research interests are in image analysis of multidimensional microscopy images, computer vision and motion analysis for cellular dynamics, and in mathematical approaches for biological imaging. He is a Fellow of the IEEE, past Chair of the IEEE SPS Bio Imaging and Signal Processing Technical Committee (BISP-TC), a senior area editor of the IEEE Signal Processing Letters, and a member of the Editorial Board of the journals Medical Image Analysis and BMC Bioinformatics. He was the general chair of the IEEE International Symposium on Biomedical Imaging in 2008.

# Professor Roger Gunn

# Imperial College London

## Advances in the Analysis of Molecular Neuroimaging Data

Roger Gunn (RG) is Chief Scientific Officer and Head of Analysis at Imanova, Professor of Molecular Neuroimaging in the Division of Brain Sciences at Imperial College London and Visiting Professor in the Department of Engineering at the University of Oxford. RG is an Applied Mathematician who received a PhD in bio-mathematical modelling of Positron Emission Tomography (PET) data from the University of Warwick in 1996. RG has worked on the development of quantitative mathematical techniques for the analysis of PET and MRI data and their application in studies of pathophysiology and drug development. RG has worked previously at the MRC Cyclotron Unit (UK), Montreal Neurological Institute (Canada) and GSK. In October 2011 he took up the position of Chief Scientific Officer at Imanova, which is a joint venture between the three major London universities (Imperial, UCL, Kings) and the MRC targeted at providing R&D imaging capabilities to the London Universities and Commercial organisations.

His research interests are focused on the application of mathematical biology and molecular imaging techniques to disease understanding and drug development in the neurosciences. RG has led the development of bio-mathematical tools for in silico screening of candidate imaging probes, quantitative analysis of dynamic PET data, measurement of BBB transport of drugs, PK/PD modelling of drug-target interactions, adaptive imaging study designs and characterization of novel PET tracers for a wide range of neurochemical targets including dopamine, serotonin, histamine, glycine, opioids, NK1, phosphodiesterase 4 and 10. RG has published over 120 peer reviewed journal articles in the field of imaging and has served on organising and programme committees at numerous conferences - he was recently Chair of BrainPET 2011 and organizer of the International PET Pharmacokinetics courses in 2010 & 2011.

# Novel Imaging Methods and Applications

# Automatic Generation of Synthetic Retinal Fundus Images

Samuele Fiorini[12]
samuele.fiorini@studenti.unipd.it
Lucia Ballerini[2]
luciaballerini@computing.dundee.ac.uk
Emanuele Trucco[2]
manueltrucco@computing.dundee.ac.uk
Alfredo Ruggeri[1]
alfredo.ruggeri@unipd.it

[1] Department of Information Engineering
University of Padova
Padova, ITALY

[2] VAMPIRE project, CVIP group
School of Computing
University of Dundee
Dundee, UK

**Abstract**

This study aims to generate *synthetic* and *realistic* retinal fundus colour images, similar in characteristics to a given dataset, as well as the values of all morphological parameters. A representative task could be, for example, the synthesis of a retinal image with the corresponding vessel tree and optic nerve head binary map, measurement of vessel width in any position, fovea localisation and so on. The presented paper mainly focuses on the generation of non-vascular regions (i.e. retinal background, fovea and optic disc) and it is complemented by a parallel study on the generation of structure and texture of the vessel network. To synthesise convincing retinal backgrounds and foveae, a *patch-based* algorithm has been developed; *model-based* texture synthesis techniques have also been implemented for the generation of realistic optic discs. The validity of our synthetic retinal images has been demonstrated by *visual inspection* and *quantitative experiments*.

## 1 Introduction

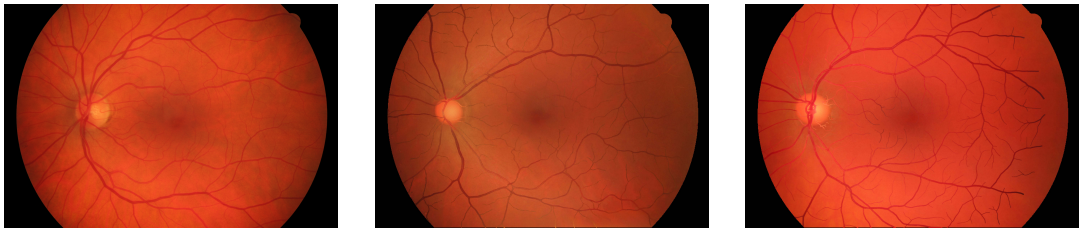
Retinal fundus imaging is a very helpful tool for the diagnosis of many diseases such as diabetes, glaucoma and cardiovascular conditions [1]. This is reflected by the large number of algorithms for optic disc (OD) detection, vessel segmentation and width measurements [8]. In order to evaluate performance, limitations and clinical applicability of a medical image analysis algorithm, a validation step is mandatory. Validation can be defined as *the process of showing that an algorithm performs correctly by comparing its output with a reference standard* [11]. Algorithm validation on manually annotated ground truth (GT) images is a well documented and necessary practice. However, it is also well known that obtaining large quantities of annotated medical GT images is an expensive and laborious task.

The main purpose of this work is the generation of synthetic high resolution retinal fundus images (Figure 1a) in a controlled manner along with a synthetic GT free from inter-observer variability. In the synthesised retinal phantoms, textural anatomical features can be modified to simulate a wide range of parameters and different populations. Our synthetic data would allow extensive validation of algorithms for segmentation and analysis of

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

anatomical structures. The final goal is the public release of a new database containing synthetic retinal images including OD and macula centre coordinates, vessel binary maps and width at any position and artery/vein classification. All these measurements are of crucial importance during the validation of retinal image analysis algorithms. Medical phantom images are extensively used in many medical imaging environment, but, to our best knowledge, no previous work on the automatic generation of retinal phantoms has ever been proposed.

This initial project has been based on the publicly available *High-Resolution Fundus* (HRF[1]) *Image Database* [7], which has 45 high-quality images at the resolution of $3504 \times 2336$ pixels. The generated retinal phantoms have the same size, but images of every resolution can be synthesised by the presented method, given an appropriate dataset.



(a) (b) (c)

Figure 1: A comparison between a real retinal fundus image from the HRF dataset (a) and two synthetic phantoms generated with our method (b) and (c).

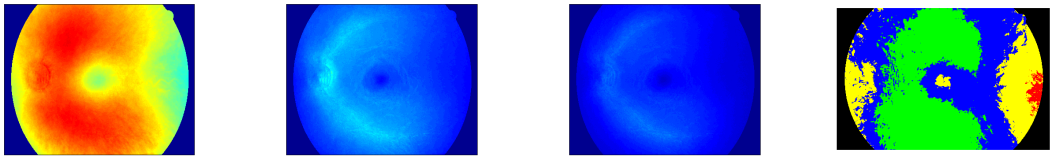
This paper is organised as follows, in Section 2 we describe the proposed method providing details for both the patch-based (Subsection 2.2) and the model-based (Subsection 2.3) texture synthesis algorithm we propose for the generation of background, fovea and OD. Section 3 describes the conducted experiments and the achieved results. Finally in Section 4 we present our conclusions and some hints for future works.

# 2 Method

## 2.1 Overview

In order to generate convincing retinal fundus image phantoms, we started from the observation of real images. At a first glance, retinal images are showing three obvious elements surrounded by an orange-red background: Optic Disc, Vessel Network and Fovea. The background presents a variety of textures. Brightness is also non-uniform across the image, usually brighter in vascular regions and darker in the fovea and in the image periphery. Intuitively the average colour distribution in retinal images could remind us of a *horseshoe* (see Figure 2). Choroidal vessels are also sometimes visible in transparency, this effect creates a real texture, very different from a mere smooth colour fill. In order to synthesise a new background preserving both non-uniform colour distribution and texture characteristics, we developed a patch-based tiling algorithm inspired by Efros and Freeman's *Image Quilting* technique [3], where novel images are synthesised by stitching together small examples of existing images. On the other hand, for the generation of synthetic ODs a novel model-based approach has been followed. This method relies on the idea of learning the distributions of key morphometric quantities from real images and realistically reproduce them.

[1]The HRF database can be downloaded at http://www5.cs.fau.de/research/data/fundus-images/



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

Figure 2: An example of colour intensity distribution map (CM) in red (a), green (b) and blue (c) channels and the correspondent clustering map (c-CM) (d).

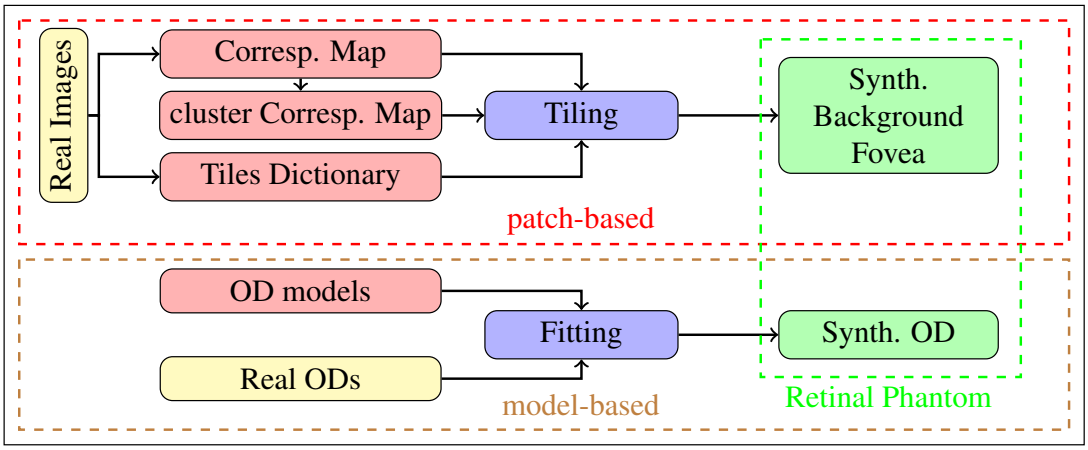



Figure 3: Block diagram depicting the overall method to generate synthetic retinal images.

## 2.2 Background and Fovea generation

In this section we will discuss about the generation of background and fovea achieved by the previously described patch-based algorithm. As we can see (Figure 3) the tiling algorithm needs three inputs: *a*) a Correspondence Map (CM); *b*) a consistent cluster-CM (c-CM) and *c*) a suitable Tiles Dictionary (TD). For our purposes, a CM is a spatial map of the corresponding colour intensity distribution over the real and the synthetic images. The CM allows our tiling algorithm to recreate a realistic background texture. Different CMs have been created by weighted averaging of 15 real preprocessed images (the healthy HRF subset). The preprocessing consisted in a naive co-registration step followed by a more sophisticated foreground (OD and vessel tree) inpainting. The co-registration step consists in a rigid rotation and a translation of all the images in order to make them share a common polar coordinate system centred on the OD. For the foreground inpainting step we took advantage of the technique proposed by Criminisi *et al*. in [2]. An example of CM is shown in Figure 2.

The tiling algorithm can generate realistic backgrounds and foveae stitching together blocks (tiles) collected from real images. For the generation of our retinal phantoms we collected more than $300,000$ vessel-free tiles $7 \times 7$ pixels and we organised them in a TD. To facilitate the research of matching blocks performed by the proposed tiling algorithm, all the collected vessel-free tiles have been grouped into clusters using a pixel-wise K-means algorithm. The number of clusters (4) has been optimised using the Akaike Information Criterion [6] and the feature selected for the clustering were the red and green intensity value of each pixel. Each tile has then been assigned to a unique cluster by means of majority vote. To link this idea with the CM, a consistent c-CM, which shows the distribution of clusters in the CM, has been evaluated as well.

For the description of the proposed tiling algorithm a syntax consistent with [3] has been chosen. To assess the colour similarity among the tiles, the original *error surface* $e$ has been replaced with a *new error surface* $\varepsilon$ defined in Equation 1 (where the symbol $\circ$ indicates the element-wise matrix product and the two indexes $n_{<C_1,C_2>}$ and $\gamma_{<C_1,C_2>}$ are, respectively, the grey level and the chroma similarity indexes described in [5]).

$$\varepsilon_{<C_1,C_2>} = 1 - n_{<C_1,C_2>} \circ \gamma_{<C_1,C_2>} \tag{1}$$

The tiling algorithm workflow can be described as follows. Let $\Phi$ be the desired output image and $\tau_k(i)$ the *i-th* element of the TD assigned to the cluster $\omega_k$. The final result can be achieved going trough $\Phi$ in raster scan order in steps of the size of the tiles minus an overlap (3 pixels) and iterating the following steps:

1. Randomly pick up successive $\tau_k(i)$ from $\omega_k$ (where $k$ is defined by the c-CM) until the best match, defined as the tile that locally satisfies the two constraints below, is found:
   - local consistence with the CM: the Euclidean distance between the picked-up tile and the correspondent tile in the CM is less than a threshold $T_1$;
   - local consistence among the neighbouring blocks in the overlap areas: more than the 90% of the pixels values of the evaluated error surface $\varepsilon$ must be less than a threshold $T_2$.
2. Compute the Cumulative Minimum Error between the overlapping areas defined as $E_{i,j} = \varepsilon_{i,j} + min(E_{i-1,j-1}, E_{i-1,j}, E_{i-1,j+1})$ where $i = 2,\ldots,N$ $j = 2,\ldots,M$ and $\varepsilon$ is $N \times M$.
3. Finally, place the new tile cutting out its edges as delineated by a Best Boundary Cut (i.e. the minimum value in each row/column of $E$) and repeat.

The two thresholds ($T_1$ and $T_2$) have been heuristically determined (10 and 0.05 respectively), as well as the default tiles and overlap dimension.

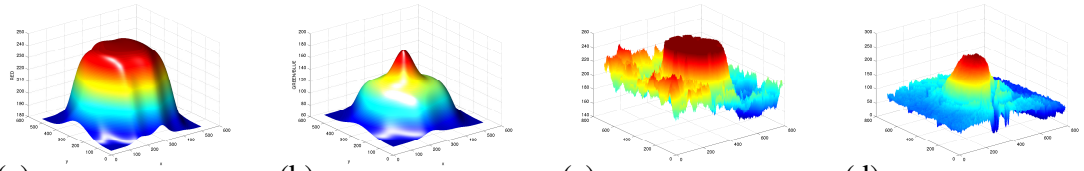
## 2.3 Optic Disc generation

The OD appears usually as a slightly irregular ellipse with a small inner depression, called cup. The edge of the OD is more evident in the red channel, while the inner cup is usually best highlighted in the green and blue ones. Therefore, to reproduce the same visual effect, we developed two parametric intensity models, one for the red (Equation 2) and one for both the green and the blue channels (Equation 3, where the meaning of the two parameter vectors $\bar{p} = [x_0, y_0, z_0, a, \sigma_1, A, \omega]$ and $\bar{q} = [\bar{p}, x_0, y_0, \sigma_2, k]$ will be clarified below):

$$f_R(x,y,\bar{p}) = z_0 - \frac{1}{a + \exp\left[-\left(\frac{x-x_0+A\cos(\omega t)}{\sigma_1}\right)^2 - \left(\frac{y-y_0+A\cos(\omega t)}{\sigma_1}\right)^2\right]} \tag{2}$$

$$f_{GB}(x,y,\bar{q}) = f_R(x,y,\bar{p}) + k\exp\left[-\left(\frac{x-x_0}{\sigma_2}\right)^2 - \left(\frac{y-y_0}{\sigma_2}\right)^2\right] \tag{3}$$

The two models refer to a three-dimensional space, where the horizontal plane $(x,y)$ is the image plane and the vertical axis $z$ is the colour intensity expressed in the $(r,g,b)$ colour



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

Figure 4: OD intensity models on the red (a), the green and the blue (b) channels. Real OD data on the red (c) and green (d) channels.

space. The OD generation workflow is described in the lower section of Figure 3. The parameters of the models have been estimated over 30 HRF images (15 healthy and 15 diabetic retinopathy) using a weighted nonlinear least squares criterion. Looking at Figure 4 an explanation of the meaning of the parameters could be provided as follows. While $(x_0, y_0, z_0)$ simply control the translation of the surface along the axes, $\sigma_1$ controls the spread of the surface in the horizontal plane (similarly to the standard deviation in the Gaussian component, $\sigma_2$). Tuning $a$ we can also modulate the amplitude of the surface. Figure 4a has a low frequency oscillating sinusoidal term that models the irregularity on the OD edges (the cosine in Equation 2 is evaluated in $[0, \cdots, 2\pi]$). Equation 3 presents a symmetrical Gaussian surface that models the inner cup pallor. A suitable probability distribution has been evaluated for each parameter. Convincing ODs can be then synthesised by random sampling a set of parameters from the relative estimated distribution and then evaluating with them Equation 2 and 3.

# 3 Results

This paper focuses on the generation of non-vascular regions; our preliminary retinal phantoms present synthetic backgrounds and ODs, but vessel networks collected from real healthy HRF images using the provided manually segmented vessel map. Figure 1 shows a comparison between a real retinal fundus image (a) and two examples from our synthetic phantoms databases (b,c). The first results look promising: the generated retinal backgrounds seem to reproduce a realistic colour intensity distribution and a realistic representation of the fovea has been achieved as well. The synthetic ODs look also realistic, they show a natural colour intensity with a correctly placed inner cup and their edge looks realistically merged with the surrounding background.

The main intended purpose of our synthetic image generator is to provide a dataset along with GT images (all parameters known) to validate retinal image analysis algorithms. Assuming that the images are affected by additive Gaussian noise, we compared the estimation of the noise variance $\sigma^2$ (evaluated as suggested in [4]) between 15 real HRF and 15 synthetic images. The obtained values are $\sigma^2_{real} = 0.020$ and $\sigma^2_{synth} = 0.022$. Then, we validated our results proving that in the non-vascular regions no vessel-like patterns are created by our method. No intensity adjustment has been provided for the vessel network, so the considered test only concerns the false positive values. We used the well-known algorithm by Soares *et al*. [9] to detect vasculature maps on 15 retinal phantoms and on 15 healthy HRF images provided with manual GT; then we compared the two obtained false positive rate ($FPR = \frac{FP}{FP+TN}$). Using a two-sample Komolgorov-Smirnov test no significant difference at the 5% level can be observed between the FPR obtained from the vessel detection on the two datasets. The average FPR is 0.0097 ad 0.0112 on real and on synthetic images, respectively.

# 4 Conclusions

In this paper a novel technique for the automatic generation of synthetic retinal fundus images has been presented. A realistic representation of the non-vascular regions (background, fovea and OD) has been achieved and we are currently working on the generation of structure and texture of synthetic vessel trees as well as retinal abnormalities. To our best knowledge, no similar method has been reported in literature. The validity of our retinal phantoms has been supported by visual inspection and quantitative experiments. We are aware of the preliminary nature of the conducted test and we plan a more complete analysis with automatic estimates of further parameters using VAMPIRE [10] and subjective evaluations by clinicians. However, the presented results let us sense that we are correctly moving towards the generation of a synthetic dataset for the validation of retinal image analysis algorithms.

**Acknoledgements:** Research partially supported by Leverhume trust grant RPG-419. We thank Roberto Annunziata and Enrico Pellegrini for valuable comments and inspiring ideas.

# References

[1] M. D. Abràmoff, M. K. Garvin, and M. Sonka. Retinal imaging and image analysis. *IEEE T-MI*, 3:169–208, 2010.

[2] A. Criminisi, P. Pérez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *IEEE Trans. on Image Processing*, 13(9):1200–1212, 2004.

[3] A. A. Efros and W. T. Freeman. Image quilting for texture synthesis and transfer. In *Proc. of SIGGRAPH'01*, pages 341–346. ACM, 2001.

[4] D. Garcia. Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics & Data Analysis*, 54(4):1167–1178, 2010.

[5] P. Guobing, X. Fang, and C. Jiaoliao. A novel algorithm for color similarity measurement and the application for bleeding detection in WCE. *IJIGSP*, 3(5):1, 2011.

[6] C. D. Manning, P. Raghavan, and H. Schütze. *Introduction to information retrieval*, volume 1. Cambridge university press Cambridge, 2008.

[7] J. Odstrcilik, R. Kolar, A. Budai, et al. Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. *IET Image Processing*, 7(4):373–383, 2013.

[8] N. Patton, T. M. Aslam, T. MacGillivray, et al. Retinal image analysis: concepts, applications and potential. *Progress in retinal and eye research*, 25(1):99–127, 2006.

[9] J.V.B. Soares, J.J.G. Leandro, R.M. Cesar, et al. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. *IEEE T-MI*, 25(9):1214–1222, 2006.

[10] E. Trucco, L. Ballerini, D. Relan, et al. Novel VAMPIRE algorithms for quantitative analysis of the retinal vasculature. In *Proc. IEEE ISSNIP/BRC*, pages 1–4, 2013.

[11] E. Trucco, A. Ruggeri, T. Karnowski, et al. Validating retinal fundus image analysis algorithms: Issues and a proposal. *Investigative Ophthalmology & Visual Science*, 54 (5):3546–3559, 2013.

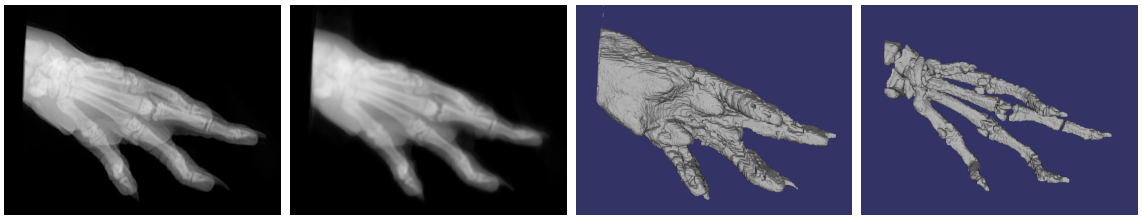
# Combined Visible and X-Ray 3D Imaging

Julien Pansiot
julien.pansiot@inria.fr
Lionel Reveret
lionel.reveret@inria.fr
Edmond Boyer
edmond.boyer@inria.fr

INRIA Grenoble Rhône-Alpes
655 Avenue de l'Europe,
38334 Saint Ismier cedex
France

**Abstract**

This paper considers 3D imaging of moving objects and introduces a technique that exploits visible and x-ray images to recover dense 3D models. While recent methods such as tomography from cone-beam x-ray can advantageously replace more expensive and higher-dose CT scanners, they still require specific equipment and immobilised patients. We investigate an alternative strategy that combines a single x-ray source and a set of colour cameras to capture rigidly moving samples. The colour cameras allow for coarse marklerless motion tracking, which is further refined with the x-ray information. Once the sample poses are correctly estimated, a dense 3D attenuation model is reconstructed from the set of x-ray frames. Preliminary results on simulated data compared to ground-truth as well as actual in-vivo experiments are presented.
**Keywords:** motion capture; tomography; x-ray; colour video.



Ground truth (256 slices) | Reconstruction (from 32 views) | Skin isosurface | Skeleton isosurface

Figure 1: Rat paw volumetric reconstruction from a sparse set of simulated visible and x-ray images (32 frames from a rotational motion). Left-to-right: ground truth CT volume, volumetric reconstruction from 32 views rendered as planar x-ray, isosurface for soft tissues interface, and skeleton interface.

## 1 Introduction and related work

The capture of movements to study functional analysis has gained great attention in the recent years with the improvement of acquisition systems. In that respect, the ability to capture simultaneously the motion of both internal structures such as the skeleton and external surfaces is a challenge with promising applications. Whilst a range of solutions are available

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

for motion capture and analysis in 3 dimensions using visual cues, most are solely limited to the recovery of surface information. Conversely, radiography allows for the capture of inner structures in motion, but is mostly restricted to 2 dimensions. In this paper we proceed one step towards full volumetric 3D motion capture by investigating the case of rigidly moving samples, relying on a combination of visible and x-ray cameras, and assuming no prior knowledge on the captured sample.

Our aim is to recover dense 3D models of moving objects. The existing two main strategies for radiography are the regular planar x-ray and the Computed Tomography (CT) scanner. While the latter produces high-resolution dense 3D models, it suffers three main drawbacks: higher radiation dose for the patient [10], higher cost and, more importantly here, immobility of the patient during the procedure. For these reasons, specific procedures have been developed to generate volumetric attenuation models from a limited number of regular cone-beam x-ray sources, such as full 3D models from isocentric/orbital C-arm [10], and increasingly well resolved models for breast tomosynthesis [11]. Model-based methods using bi-planar x-ray beams [1] even allow motion capture [2], however such methods require strong prior models and usually manual intervention.

In order to compute a dense 3D attenuation model while limiting the patient dose, allowing for motion capture and incidentally reducing cost, we propose a novel approach that combines regular colour cameras with standard x-ray imagery to acquire a rigidly but freely moving markerless sample. Our framework keeps all equipment static, eliminating the need for specific isocentric C-arms or other tightly controlled systems [10]. The approach does not use prior anatomic models either, hence this is neither a registration problem [11], nor a model fitting issue [1] and it allows for unknown shapes. The paradigm we follow here does not consider motion as noise but, on the contrary, as one key to 3D reconstruction.

The closest approach to ours has been proposed by Sidky *et al*. [9] where a reconstruction method for any type of calibrated x-ray imaging device under the assumption of limited angles and/or number of views is presented. Schumacher *et al*. [8] also proposed a method that combines reconstruction and motion correction in SPECT imaging, without requiring colour cameras. This method is however tailored for motion correction and hence appears to cater only for a limited range of motion. The use of colour video for motion detection during e.g. SPECT has been covered for example by [6], but requires markers. A markerless method has been proposed by [4], but is again targeted to small motion correction.

To the best of our knowledge, our approach is the first to use motion, through the combination of visible and x-ray images, to perform 3D radiography. Potential medical applications include affordable dental 3D imaging, 3D imaging in confined environments, and, when non-rigidities will be handled, joints inner dynamics.

The remainder of this paper is composed of the following: in Section 2 we describe our reconstruction method, in Section 3 we present our results from both simulated and in-vivo data, before concluding in Section 4.

## 2 Proposed method

The proposed method relies on a single x-ray source combined with a set of regular colour cameras. The colour cameras are used to build a 3D surface model that is tracked over time. The motion estimation of this model is refined using the x-ray images. Once the motion is known, a dense 3D attenuation model can be reconstructed from the x-ray data. The overall pipeline is illustrated in Figure 2.

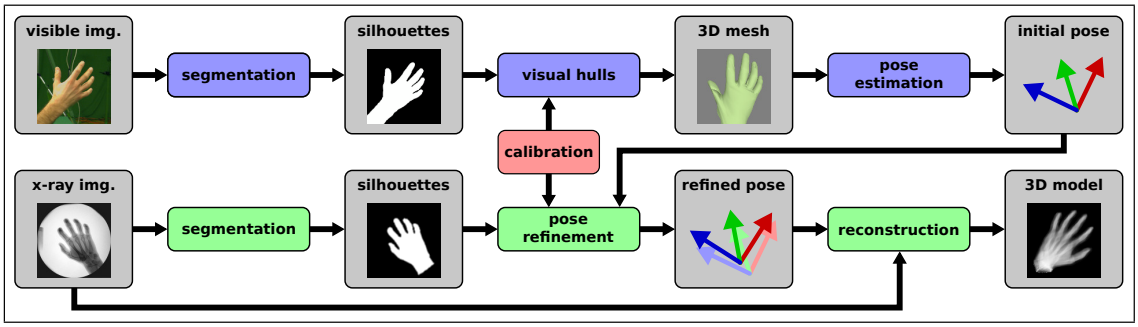



Figure 2: Overview of the proposed pipeline for combined visible and x-ray 3D imaging.

## 2.1 Calibration

Both visible and x-ray cameras are fully calibrated in world coordinate system prior to acquisition. This operation is performed in two stages to minimise the x-ray usage. Firstly, the eight video cameras are fully calibrated (intrinsic and pose) on their own, using a moving LED wand within a single bundle adjustment framework. Secondly, a lightly coloured lead ball is displaced within the entire capture volume, triangulated by the visible cameras, before the x-ray camera is finally fully calibrated knowing the 3D from calibrated video cameras.

## 2.2 Initial pose estimation

Since the sample can move freely within the capture volume, it is necessary to estimate its pose prior to reconstruction. Using the colour cameras only, the silhouettes of the sample segmented from the background are processed with a polyhedral visual hull algorithm [3] to generate a 3D mesh of the sample per frame (see fig. 2). The remaining meshes of the sequence are then registered to a reference pose model, e.g. the first frame, with a robust Iterative Closest Point (ICP) implementation with outlier detection. This provides an initial pose, i.e. rotation and translation, estimation for each frame and with respect to the reference frame.

## 2.3 Pose refinement

Since the visual hull meshes exhibit artefacts due to the limited number of colour cameras, the initial pose estimation is corrupted. To improve the pose, we exploit directly the x-ray silhouettes, segmented from the x-ray background light.

To this aim, we rely on the assumption that if the poses were perfectly registered, the volume reprojection on the x-ray plane should equate perfectly to the x-ray silhouettes. For simplicity, we consider the model to be static and captured by moving cameras, with intrinsic parameters from the initial calibration, and extrinsic from the initial pose estimation. A cost function penalising differences between the original x-ray silhouettes and the reprojected model is used within a gradient descent framework to refine the poses iteratively. If $C$ is the set of $N$ camera parameters, $I$ the set of $N$ segmented images, $M$ a mesh, 'proj' the camera projection operator, and 'vh' the visual hull operator, we try to estimate:

$$\arg\min_{C} \left( \sum_{i=1}^{N} |\mathrm{proj}_{C_i}(M) - I_i| \right) \quad st. \quad M = \mathrm{vh}_C(I). \tag{1}$$

This extra stage also reduces slight spatial and temporal misalignment between the visible and x-ray systems, providing more robust results.

### 2.4 Reconstruction

Once the pose is known, the Kaczmarz method [5], also referred to as Algebraic Reconstruction Technique (ART) is used to reconstruct iteratively the volumetric attenuation. For enhanced performance, the views are ordered such that two successive points of view are most orthogonal, and the reconstruction takes places only within the x-ray visual hull.

Given the volume $X$ being reconstructed, the observed images $A$ with a saturated attenuation $A_{sat}$, and a relaxation parameter $\lambda_i$, we exploit the projection matrix sparseness by precomputing each camera ray defined as the list of the $N$ intersected voxels in the reconstructed volume $p$ and their associated weights $w$. For each ray $R_i = \{(p_{i,1}, w_{i,1}), ..., (p_{i,N}, w_{i,N})\}$ associated to the image pixel value $A_i$, we then compute the norm, the current integrated attenuation on the ray $x_i$, and the weighted difference with the observed pixel value, $d_i$:

$$\|R_i\|^2 = \sum_{j=1}^{N} w_{i,j}^2, \qquad x_i = \sum_{j=1}^{N} w_{i,j} X\left[p_{i,j}\right], \qquad d_i = \lambda_i \frac{A_i - x_i}{\|R_i\|^2}. \tag{2}$$

The voxel attenuations are then updated as follows:

$$X^{n+1}\left[p_{i,j}\right] = \begin{cases} 0 & \text{if } X^n\left[p_{i,j}\right] + d_i w_{i,j} < 0 \text{ (positivity constraint)}, \\ X^n\left[p_{i,j}\right] & \text{if } d_i < 0 \text{ and } A_i = A_{sat} \text{ (prevent saturation artefacts)}, \\ X^n\left[p_{i,j}\right] + d_i w_{i,j} & \text{in other cases.} \end{cases} \tag{3}$$

The relaxation parameter $\lambda_i$ is weighted for each pose based on the leave-one-out score calculated as in eq. (1) in order to favour poses which appear more reliably estimated.

Since the amount and nature of the observed data may render the problem locally ill-posed, high-frequency noise is observed in the results. Using the assumption that living organisms can be modeled by a set of relatively homogeneous tissues, a 3D adaptation of Rudin *et al.* seminal non-linear Total Variation (TV) [7] is enforced on the model in between iterations for regularisation, as proposed by [9]. This improves dramatically the results, as illustrated in Figure 3.

## 3 Experiments and results

### 3.1 Simulated data

The combined x-ray and visible images can be simulated by rendering real volumetric CT data using raycasting. In this experiment, we rendered the CT model of a rat paw ($256^3$ voxels), simulating a system of eight colour cameras and one x-ray device. The combined set of 2D images was then processed through the whole pipeline. The effective projection resolution for both systems was $600 \times 600$ pixels. The model was artificially moved continuously for 32 frames, following roughly a 180-degree rotation. This allowed for comparison with ground-truth data. Figures 1 and 3 show the original model as well as the reconstructed volume.

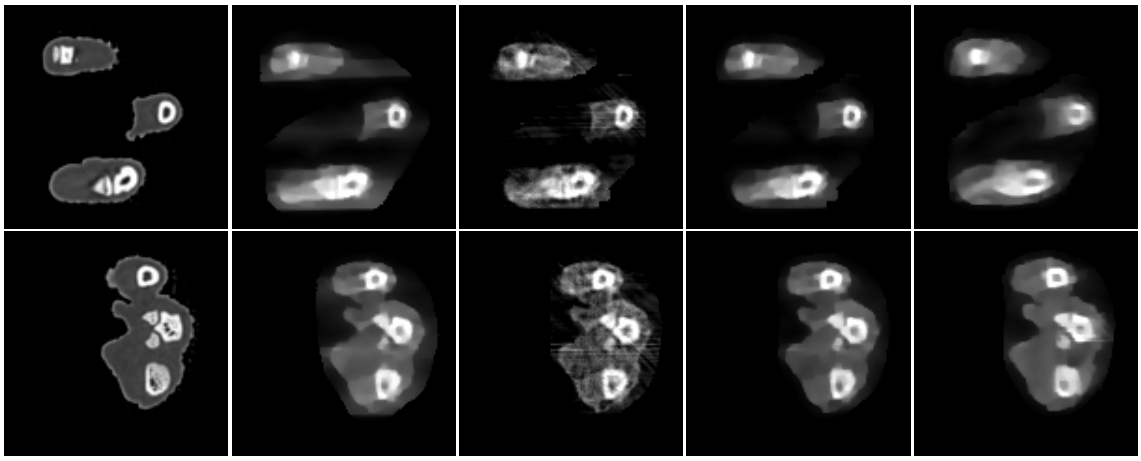



Figure 3: Two rat paw reconstructed slices from simulated visible and x-ray images (32 frames). Left-to-right: ground-truth CT data (256 slices), motion estimated with ICP and TV regularisation, motion refined but without TV, motion refined and TV, true motion and TV. Slight motion estimation errors are clearly corrected by the refinement stage, but not as well as with true motion. TV regularisation reduces significantly the high-frequency noise.

## 3.2 In-vivo experiment

A system composed of eight colour cameras (2048×2048 pixels) and one Siemens ARCADIS Avantic x-ray C-arm (1024×1024 pixels) was deployed for this preliminary experiment. The reconstruction volume ($256^3$ voxels) covered a cube of 20×20×20 cm, *ie.* a resolution of 0.78mm/voxel.

Three volunteers (one of them shown here) were asked to perform roughly a 180-degree wrist circumduction with fingers kept as rigid as possible within the working volume, during which 32 frames were captured. The colour cameras recorded at 30 fps, whilst the x-ray acquired at 10 fps to minimise the dose, both being only coarsely synchronised. Figure 4 illustrates the preliminary results of the proposed method.

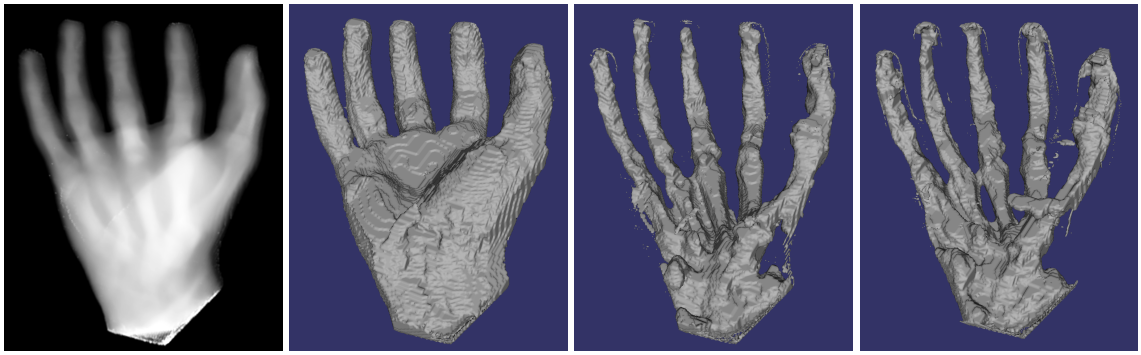


Figure 4: Human hand reconstruction with a combined visible & x-ray system (32 frames). Left-to-right: volume rendered as planar x-ray, soft tissues thresholding, skeleton thresholding, skeleton thresholding without pose refinement: some phalanges are thicker, the distal noisier. The soft tissues exhibit a concavity that cannot be modelled solely by visual hulls.

It is evident from the results that whilst the simulated data demonstrates the proposed

approach's great potential, the preliminary reconstruction of in-vivo data is noisier. We attribute this difference to the fast shutter noise in x-ray images (in particular in the metacarpal area) as well as non strictly rigid motion (on the finger tips and the thumb in particular). The former should be improved by carefully planned x-ray acquisition parameters.

# 4 Conclusion and future work

In this paper we have shown that a multi-camera system can advantageously complement a single static x-ray imaging device and enable 3D motion imaging of rigidly moving samples. Future work includes the analysis of the achievable quality based on the recorded motion and non-rigid motion handling using rigid-by-part modelling.

# Acknowledgments

This project was funded by KINOVIS (ANR-11-EQPX-0024) and MORPHO (ANR-10-BLAN-0206), with help from Amiqual4Home (ANR-11-EQPX-0002).

# References

[1] S. Benameur, M. Mignotte, H. Labelle, and J. A. De Guise. A hierarchical statistical modeling approach for the unsupervised 3-D biplanar reconstruction of the scoliotic spine. *Biomedical Engineering, IEEE Transactions on*, 52(12):2041–2057, 2005.

[2] E. L. Brainerd, D. B. Baier, S. M. Gatesy, T. L. Hedrick, K. A. Metzger, S. L. Gilbert, and J. J. Crisco. X-ray reconstruction of moving morphology (XROMM): precision, accuracy and applications in comparative biomechanics research. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 313 (5):262–279, 2010.

[3] J.-S. Franco and E. Boyer. Exact polyhedral visual hulls. In *BMVC*, pages 329–338, 2003.

[4] B. F. Hutton, A. Z. Kyme, Y. H. Lau, D. W. Skerrett, and R. R. Fulton. A hybrid 3-D reconstruction/registration algorithm for correction of head motion in emission tomography. *Nuclear Science, IEEE Transactions on*, 49 (1):188–194, 2002.

[5] S. Kaczmarz. Angenäherte auflösung von systemen linearer gleichungen. *Bulletin International de l'Académie Polonaise des Sciences et des Lettres. Classe des Sciences Mathématiques et Naturelles. Série A, Sciences Mathématiques*, pages 355–357, 1937.

[6] J. E. McNamara, P. H. Pretorius, K. Johnson, J. M. Mukherjee, J. Dey, M. A. Gennert, and M. A. King. A flexible multicamera visual-tracking system for detecting and correcting motion-induced artifacts in cardiac SPECT slices. *Medical physics*, 36(5):1913–1923, 2009.

[7] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[8] H. Schumacher, J. Modersitzki, and B. Fischer. Combined reconstruction and motion correction in SPECT imaging. *Nuclear Science, IEEE Transactions on*, 56(1):73–80, 2009.

[9] E. Y. Sidky, C.-M. Kao, and X. Pan. Accurate image reconstruction from few-views and limited-angle data in divergent-beam ct. *Journal of X-ray Science and Technology*, 14(2):119–139, 2006.

[10] J. H. Siewerdsen, D. J. Moseley, S. Burch, S. K. Bisland, A. Bogaards, B. C. Wilson, and D. A. Jaffray. Volume CT with a flat-panel detector on a mobile, isocentric C-arm: pre-clinical investigation in guidance of minimally invasive surgery. *Medical physics*, 32(1):241–254, 2005.

[11] G. Yang, J. H. Hipwell, D. J. Hawkes, and S. R. Arridge. A nonlinear least squares method for solving the joint reconstruction and registration problem in digital breast tomosynthesis. In *MIUA*, pages 87–92, 2012.

# Photometric Stereo Reconstruction for Surface Analysis of Mucosal Tissue

Khemraj Emrith[1]

Greg Slabaugh[2]

Andrew Poullis[3]

Christopher Groves[3]

Melvyn Smith[1]

[1] University of the West of England
Bristol, UK

[2] City University London
London, UK

[3] St. George's Hospital
London, UK

**Abstract**

We introduce photometric stereo reconstruction for surface analysis of mucosal tissue. Photometric stereo captures images of a surface from a single camera, but illuminated from different light sources. From the set of images, a detailed surface reconstruction is estimated. We apply photometric stereo principles to mucosal tissue, which raises technical challenges due to its specular appearance. Following reconstruction of the surface, we apply surface analysis techniques based on differential geometry to detect regions that are locally spherical, suggestive of polyps. The research described in this paper is a first step towards a real-time clinical support system based on 3D reconstruction of human tissue for gastereoentrologic applications such as colonoscopy.

## 1 Introduction

Photometric stereo [8] is a computer vision technique that estimates a dense set of normal vectors of a surface photographed with a fixed camera and multiple light sources. Under the assumption of Lambertian reflectance, and a minimum of three known light source locations, surface normal vectors can be determined and then integrated to reconstruct a detailed surface topography in three dimensions.

In the context of medical imaging, photometric stereo (PS) has recently been explored in applications ranging from dermatology [3] to laparoscopy [7]. In this paper, we apply photometric stereo to mucosal tissue, which is tissue that is typically coated with fluid secreted by a mucous membrane. Although mucosal tissue appears in many places in the body, such as the nostrils and eyelids, we are mostly interested in applications to the gastereointestinal (GI) system, extending from the mouth to the anus. Endoscopic applications such as colonoscopy examine the colon and rectum for pathology such as polyps, whereas upper GI clinical applications inspect the esophagus, stomach, and duodenum for conditions such as Barrett's esophagus. Due to its capability for real-time 3D reconstruction of high resolution surface details, in addition to colour images, photometric stereo has much potential for detection, diagnosis, and monitoring of disease in the GI tract.

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

However, traditional endoscopic equipment is limited to 2D imaging, using standard video or multispectral techniques like narrow-band imaging. Recently there has been been growing interest in using two-camera endoscopy to create video feeds that can be viewed using 3D glasses. However, such systems do not actually acquire a 3D surface, rather two video streams that are presented to the endoscopist. We argue that additional useful information can be extracted from a 3D reconstruction of the surface using photometric stereo, and the recovered geometry can be analysed, ultimately as part of a clinical workflow.

In related work, Lv et al. [4] apply photometric stereo to reconstruct a phantom model of the human tongue to capture subtle surface anatomy such as the papillae. Perhaps most related to our work is that of [2, 5], which applies photometric stereo to both ex-vivo and in-vivo mucosal tissue. This work demonstrates the promise of photometric stereo to differentiate normal from abnormal tissue, even at diminutive sizes, but also highlights the challenge of imaging mucosal tissue, which produces specular highlights due to the fluid on the tissue.

In this paper we also apply photometric stereo to mucosal tissue. However, to our knowledge, this is the first paper to perform 3D shape analysis of mucosal tissue reconstructed using photometric stereo. We produce a proof-of-concept results demonstrating that photometric stereo is capable of capturing important geometric details that can be analysed using computational geometry. Inspired by the CT Colonography computer-aided detection literature, we apply simple differential geometric measures to the heightmap reconstructed from photometric stereo, and use them to define regions of interest in image, the first step in a computer-aided detection algorithm to identify suspicious pathology.

## 2 3D Reconstruction of Mucosal Tissue

In this section we investigate the effect of illuminating a porcine gut wall and capture images to recover the 3D surface topography of different patches of the GI tract. A key challenge is that mucosal surfaces are moist, producing specular reflections when photographed. We describe our vision-based approach and how it addresses this challenge.

Using a porcine gut model, photometric images were captured using a six-light source photometric stereo setup. Photometric stereo assumes diffuse reflectance from the illuminated surface. However, as shown in [2, 3], a Lambertian model can be used to approximate the reflectance characteristics of the mucosal surface by discarding patches with specular reflection. In this study we use a least squares approximation approach to estimate the surface in the presence of the highlights. Several areas of the porcine gastrointestinal tract were scanned. A Gig-E Basler Pilot camera with a resolution of $2448 \times 2050$ pixels was used to capture sets of photometric images from each area investigated. Figure 1 shows one set of photometric images, containing considerable specular highlights, captured from the colon part of the gastrointestinal tract. The highlights saturate the sensor on the imaging device, making the recovery of the surface at that those locations extremely challenging. If left unhandled, highlights result in discontinuities in the surface normals that are used estimate the 3D surface topography. The images are captured in sequence from six different illumination directions for the same observation point. The intensities and the illumination data is stacked, respectively, to form an intensity vector $\mathbf{I}(\mathbf{x},\mathbf{y}) = (I_1(x,y), I_2(x,y) \ldots I_6(x,y))$ and an illumination matrix $[\mathbf{L}] = (l_1, l_2, l_3)$. Assuming Lambertian reflectance, we write

$$\mathbf{I}(x,y) = p(x,y)\,(\mathbf{L}\cdot\mathbf{n})\,, \tag{1}$$

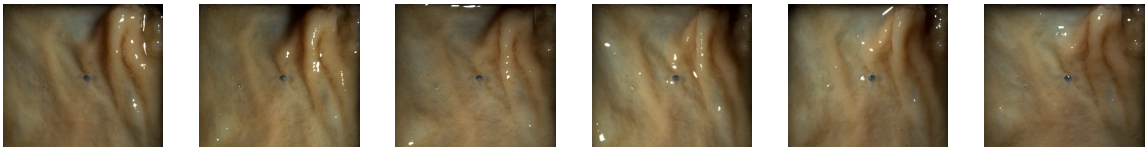
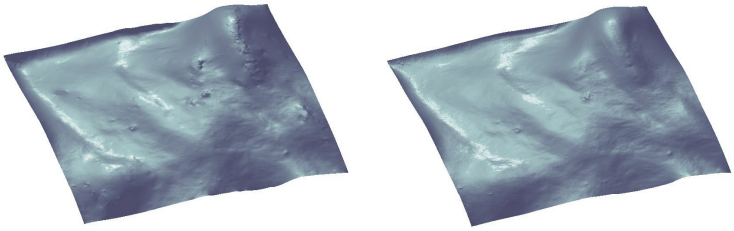


Figure 1: A sequence of six photometric images captured from a patch of duodenum of the porcine gut model used (zoom in for detail).



(a) (b)

Figure 2: Recovered 3D mucosal tissue under uniform illumination (a) without highlights elimination (b) with highlights elimination.

where $p(x,y)$ is the albedo and $\mathbf{n}$ is the surface normal. Assuming that matrix $\mathbf{L}$ is known, and is at least of rank 3 (i.e. three or more light sources are used), then the albedo and surface normal can be uniquely calculated using linear least-squares:

$$p = |\left(\mathbf{L}^T\mathbf{L}\right)^{-1}\cdot\mathbf{L}^T\cdot\mathbf{I}| \quad (2)$$

$$\mathbf{n} = \left(\left(\mathbf{L}^T\mathbf{L}\right)^{-1}\cdot\mathbf{L}^T\cdot\mathbf{I}\right)/p, \quad (3)$$

To eliminate discontinuities (highlights and shadows) that arise during capture, the stacked intensity values from the six different light sources are sorted at every pixel location and the smallest and largest values are eliminated as they are bound to represent either shadows (low intensity) or highlights (high intensity). The remaining intensity values at every pixel is then used to recover the 3D surface geometry of the patch of mucosal tissue being investigated. Figure 2(a) and (b) shows 3D surfaces from the sequence shown in Figure 1 rendered under uniform illumination. We observe that although the combined amount of highlights is not high, the recovered surface shown in Figure 2(a) contains distorted patches when the highlights are not eliminated as compared to the recovered 3D surface with highlights eliminated (see Figure 2(b)).

# 3 Surface Analysis

In this section we describe our approach to surface analysis, based on the recovered height field described in the previous section.

## 3.1 Differential geometry of height fields

A height field $f(x,y)$ describes a surface of the form $\mathbf{S} = (x,y,f(x,y))$, which provides a mapping from $\mathbb{R}^2 \rightarrow \mathbb{R}^3$.

In this paper we are interested in the differential geometric structure of the surface, which is characterised using the first and second fundamental forms [1], composed of coefficients of first order surface derivatives $E = \mathbf{S}_x \cdot \mathbf{S}_x$, $F = \mathbf{S}_x \cdot \mathbf{S}_y$, and $G = \mathbf{S}_y \cdot \mathbf{S}_y$, as well as second order surface derivatives $L = \mathbf{S}_{xx} \cdot \mathbf{n}$, $M = \mathbf{S}_{xy} \cdot \mathbf{n}$, $N = \mathbf{S}_{yy} \cdot \mathbf{n}$.

### 3.2 Surface measures

As the data is arranged on a square lattice in the form of an image, one can easily compute the above derivatives using finite differences. These derivatives are then used to compute the Mean ($H$), Gaussian ($K$), and principle curvatures ($K_1$ and $K_2$) for each point on the surface, leading the to shape index $SI$,

$$H = \frac{LN - M^2}{EG - F^2} \tag{4}$$

$$K = \frac{EN + GL - 2FM}{2(EG - F^2)} \tag{5}$$

$$K_1 = H + \sqrt{(H^2 - K)} \tag{6}$$

$$K_2 = H - \sqrt{(H^2 - K)} \tag{7}$$

$$SI = \frac{1}{2} - \frac{1}{\pi} \tan^{-1}\left(\frac{K_1 + K_2}{K_1 - K_2}\right) \tag{8}$$

The Shape Index (SI), which combines the two principal curvatures in a single measure, providing a local shape feature at each voxel. Every distinct shape, except for the plane, corresponds to a unique shape index. For example, a shape index value of 1.0 indicates a sphere-like shape (such as a polyps), while 0.75 indicates a cylinder-like shape (such as a haustral fold). Shape index can be computed at different scales by resampling the heightmap to a desired resolution.

## 4 Results

Our photometric stereo imaging system was used to produce five *ex-vivo* reconstructions of pig's gastereointestinal system, including the colon, duodenum (two images), oesophagus, and gasto-oesophageal junction, as shown in Figure 3, left three columns. A pig's gastereointestinal system is a good proxy for that of a human, with similar features such as haustral folds, and three longitudinal muscular bands (tenia). For this reason pig's colons are commonly used in training gastereoenterologists.

The heightmaps were analysed using the surface measures described in Section 3. Results are also shown in Figure 3, right two columns, for a course scale achieved by resampling the heightmap by a factor of 0.01. In this visualisation, the shape index values are clamped to the range [0.5, 1]. The values of the shape index that are highest correspond to places that are locally spherical, when both principle curvatures are large. These are good candidates for polyps, and indeed a segmentation of the shape index image may provide a useful way to establish candidate polypoidal regions, as is done in Computer-Aided Detection (CAD) systems designed for CT Colonography [9].

We perform a segmentation of the shape index using a hysteresis thresholding [6] technique, which relies on two thresholds ($T_l$, $T_h$) to define a region of interest. First, the shape index image is thresholded with the lower threshold, and connected component analysis is performed to identify connected regions with shape index above the lower threshold, $T_l$. These regions however may include anatomical regions that are not highly spherical. Each connected region is then checked based on $T_h$, and only the regions that contain at least one pixel above $T_h$ are retained. In Figure 4 we show example segmentations of the shape index overlaid on top of the original images, for $[T_l, T_h] = [0.8, 0.98]$. The region of interest

does an excellent job of segmenting an artificial polyp, with a raised, hemispherical shape, created by a gastereoentrologist and shown in the second row of the figure. The current study shows that the 3D descriptors derived are highly suitable to detect polyps which are characterised by locally spherical regions in the GI tract. An extension to this work is to capture a large dataset of mucosal tissues, excised from patients due to potential presence of polyps (benign/malignant). Such a dataset would be used to generate 3D descriptors to train an automatic polyp detection system. Classification results and validation approaches will be provided in a follow-up study.

# 5 Conclusion

We presented a technique for photometric stereo reconstruction surface analysis of mucosal tissue. Whilst the dataset in this paper is limited to five reconstructions of a porcine model, it's clear that photometric stereo has much promise in endoscopy to capture important geometric features which can be analysed using computational geometry.

This paper is a first step towards a real-time photometric stereo endoscopy system that provides clinical support to the endoscopist through surface-based image enhancement, computer-aided detection, and diagnosis. A key issue to be resolved in realising this vision is miniaturisation of the hardware to achieve a form factor that is comparable to modern video endoscopic equipment. Related to this is the challenge of performing photometric stereo in a narrow, confined space, particularly when the camera view to the surface is non-fronto parallel. One these challenges are addressed, we plan to explore in-vivo human tests and build a computer-aided detection/diagnosis system to assist the gasteroenterologist.

# References

[1] M. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice Hall, ISBN: 0132125897, 1976.

[2] Durr N. J., V. Parot, G. Traverso, W. P. Puricelli, B. J. Vakoc, N. S. Nishioka, and G. Gonzalez. Imaging colonic surface topography with photometric stereo endoscopy. In *Digestive Disease Week; American Society for Gastrointestinal Endoscopy*, 2014.

[3] Z. Liu, J. Sun, L. N. Smith, M. L. Smith, and R. Warr. Distribution quantification on dermoscopy images for computer-assisted diagnosis of cutaneous melanomas. *Medical and Biological Engineering and Computing*, 50:503–513, 2012.

[4] H. Lv, Y. Cai, and S. Guo. 3d reconstruction of tongue surface based on photometric stereo. In *Proc. International Conference on Signal Processing (ICSP)*, pages 1668–1671, 2012.

[5] V. Parot, D. Lim, G. Gonzalez, G. Traverso, Nishioka N. S., Vakoc B., and N. Durr. Photometric stereo endoscopy. *Journal of Biomedical Optics*, 18(7), 2013.

[6] M. Petrou and C. Petrou. *Image Processing: The Fundamentals*. Wiley, 2010.

[7] Collins T. and A. Bartoli. 3d reconstruction in laparoscopy with close-range photometric stereo. In *Medical Image Computing and Computer Aided Intervenions (MICCAI)*, pages 634–642, 2012.

[8] R. M. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[9] H. Yoshida and J. Nappi. Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps. *Transactions on Medical Imaging*, 20(12):1261–1274, 2001.

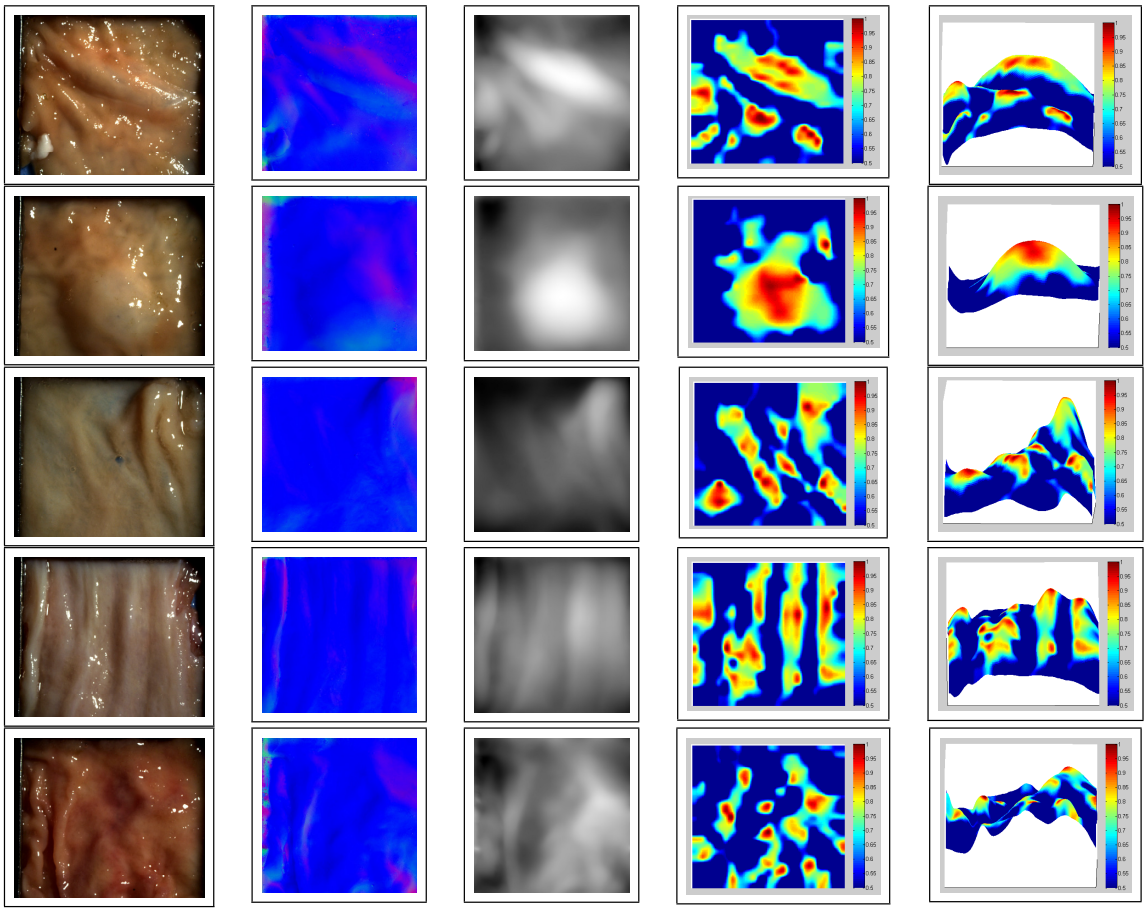
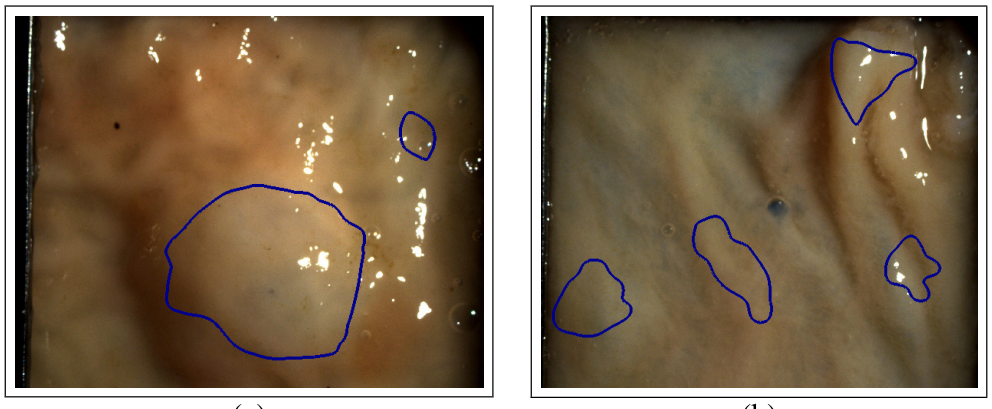


Figure 3: Data captured using the photometric stereo system. From left to right: one of the six images (left column), reconstructed normal map (second column), and reconstructed height map (middle column). From top to bottom: colon, duodenum, duodenum, oesophagus, gasto-oesophageall junction. The shape index is plotted on the surface, shown in a top view (fourth column) and side view (right column).



(a) (b)

Figure 4: Example segmentations using a hysteresis thresholding technique applied to the shape index image. These provide regions of interest that are locally spherical in the image.

# Neurological Image Analysis

# Diagnostic Plane Extraction from 3D Parametric Surface of the Fetal Cranium

Ana I. L. Namburete
ana.namburete@eng.ox.ac.uk
Richard V. Stebbing
richard.stebbing@eng.ox.ac.uk
J. Alison Noble
alison.noble@eng.ox.ac.uk

Biomedical Image Analysis Laboratory
Institute of Biomedical Engineering
Old Road Campus Research Building
University of Oxford
Oxford, UK

**Abstract**

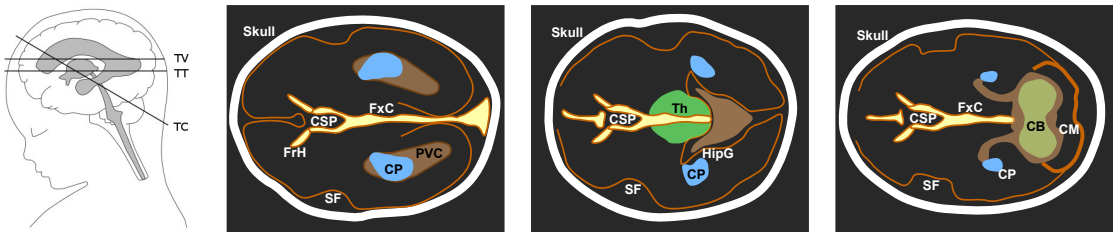
In this paper, we evaluate the viability of using a 3D parametric surface model of the fetal cranium to extract diagnostic 2D ultrasound (US) image planes and biometric measurements useful in fetal growth monitoring. The parametric surface deforms to tightly delineate the inner cranial boundary, and its topology-preserving property defines a coordinate space which allows for consistent sampling of anatomical regions. Our validation experiments comparing diagnostic planes extracted from the parametric surface to those manually selected by a clinical expert demonstrate that the planes are acceptable, and biometric measurements obtained from surface-extracted planes are within clinically-acceptable error bounds (error=4.82%, $n = 191$ fetuses).

## 1 Introduction

In obstetric care, one of the main purposes of fetal brain imaging is to identify structural abnormalities. Qualitative and quantitative image data from a fetus are compared to reference standards of normal development to enable a judgement to be made about the likelihood of an abnormality being present. Biometry of the cranium is an obvious starting point for such examinations due to its echo-bright appearance on US images. Whilst not directly measuring the brain itself, cranial size is used as a proxy for this and is incorporated into most assessments of brain structure and fetal growth. During an obstetric examination, three standard axial planes are collected from the fetal brain to assess anatomic integrity (Figure 1) [1]. However, assessment of growth from these planes is complicated by inconsistencies and subjectivity in identifying the appropriate 2D diagnostic plane from the 3D brain space, which in turn affects the skull size approximation.

In this work, we propose the use of a 3D parametrization of the fetal cranial surface to extract diagnostic planes consistently from a predetermined coordinate space to which the 3D US images are aligned. Our work benefits from the surface deformation framework described in [4] in which a B-spline spherical surface is aligned to key anatomical landmarks seen in a 3D US image of the fetal brain. Using the Levenberg-Marquardt algorithm, the surface then deforms to adhere to the inner cranial boundary through a series of discrete and continuous optimization steps [3]. During the deformation process, the surface mesh

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.




(a) Standard planes (b) TV Plane (c) TT Plane (d) TC Plane

Figure 1: **Standard Axial Planes of Fetal Head.** (a) Three axial planes used in the standard fetal neurosonography examination (from left to right): (b) transventricular (TV) plane, (c) transthalamic (TT) plane, (d) transcerebellar (TC) plane. The structures of interest in these planes are: choroid plexus (CP), posterior ventricle cavity (PVC), cavum septum pellucidum (CSP), falx cerebri (FxC), Sylvian fissure (SF), frontal horn (FrH), thalami (Th), hippocampal gyrus (HipG), cisterna magna (CM), and cerebellum (CB).

preserves topological relations between vertices, allowing for the creation of a manifold representation of the fetal skull.

The advantage of using a parametrized surface is that key points of clinical significance may be identified from which intracranial regions may be sampled and planes may be extracted on different surfaces without requiring registration between images— a challenge in fetal brain images due to underdeveloped or inconsistent neuroanatomical landmarks. In this paper, the viability of the surface at extracting the standard fetal head planes through parametrized planes whilst maintaining anatomical consistency between images of different subjects is tested, as well as the accuracy expected of the initial surface alignment provided by the user. We compare diagnostic 2D image planes selected by a clinical expert to the planes extracted from the topological surface domain, and evaluate the viability of using the surface to obtain biometric measurements useful in monitoring fetal growth.

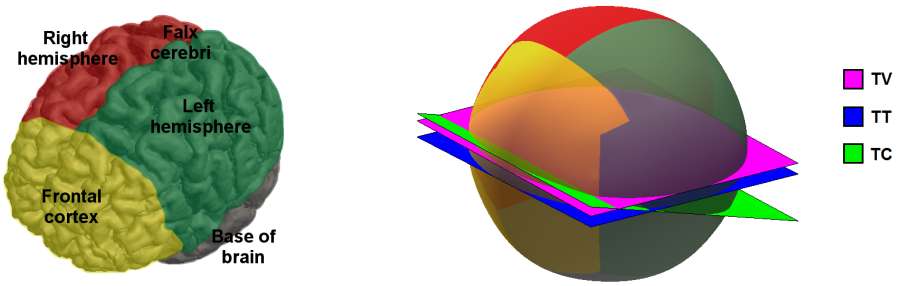
## 2 Defining Standard Planes on Cranial Domain

The standard planes were defined on the cranial surface model (Figure 2b) based on the regions they ought to intersect on the anatomically-labelled surface (Figure 2a). A single reference brain US image was used to identify the image planes, and each diagnostic plane was defined by three cranial points selected on the Euclidean image space, $C_p \in \mathbb{R}^{3\times 3}$. The points were then translated to define each standard plane on the deformed surface of the reference image, $P_j$, where $j \in \{\text{TV, TT, TC}\}$, and hence on the cranial domain.

The fetal head then underwent the surface deformation process proposed by Namburete et al. [4], and the surface points $U' \in N_U$ intersected by the standard image plane $P_j$ were used to define the standard plane $j$ on the cranial domain. Figure 2b displays the spatial definition of the standard planes on the cranial surface.

## 3 Experiments

Plane selection was performed on 52 3D US images of the fetal brain obtained from the INTERGROWTH-21$^{st}$ database, which consists of optimally healthy women with low risk of fetal abnormalities. Women included in this study ranged from $18^{+0}$ to $28^{+6}$ post-menstrual weeks' $(\text{weeks}^{+\text{days}})$ gestation. Images were collected using a Philips HD9 curvilinear probe at 2-5MHz wave frequency, and were all rendered to an isotropic voxel resolution of $0.6mm$. 2D diagnostic planes were manually-selected from the 3D brain volumes by a clinical expert,




(a) Anatomical regions (b) Planes on surface domain

Figure 2: **Standard Axial Planes on Cranial Domain.** (a) Given the anatomical regions corresponding to vertices on the cranial surface model, the standard planes can be defined on the surface (b): tranventricular (TV), transthalamic (TT), and transcerebellar (TC) planes.

which were used as the ground-truth for the planes extracted from their corresponding cranial surfaces. A further experiment compared the head circumference measurements from manually-selected ($Q$) and surface-extracted ($S$) TT planes of 191 3D US images.

## 3.1 Validation of Plane Selection

The dihedral angle between the manually-selected ($Q$) and surface-extracted ($S$) clinical planes was computed as a measure of the alignment between the planes. The dihedral angle, $\theta_d(Q,S)$, is given by:

$$\theta_d(Q,S) = \arccos\left(\frac{\mathbf{n}_Q \cdot \mathbf{n}_S}{|\mathbf{n}_Q|\,|\mathbf{n}_S|}\right) \tag{1}$$

where $\mathbf{n}_Q$ and $\mathbf{n}_S$ are the normal vectors to the manually-selected and surface-extracted planes, respectively, and the norm of a vector is defined as $|\mathbf{n}| = \sqrt{n_x^2 + n_y^2 + n_z^2}$.

It should be noted, however, that although the dihedral angle is indicative of the alignment between planes, it does not translate into a clinical index. More clinically-informative indicators are: *a*) the relationship between the anatomical structures of interest visualized in the surface-extracted plane and in the manually-selected plane, and *b*) the distance between the $Q$ and $S$ planes. For instance, a good surface-extracted TC plane would demonstrate high overlap in the cerebellar region visualized in the manually-selected plane and small average distance (in mm) from the manually-selected plane. Anatomical structure overlap is an appropriate measure because the ultrasonographer selects a suitable plane from the 3D brain space on the basis of presence, shape, and size of anatomical structures of interest. The overlap between the planes was computed using the Dice Coefficient of the manual segmentations of the following anatomical structures: choroid plexus (CP), posterior ventricle cavity (PVC), cavum septum pellucidum (CSP), and cerebellum (CB). The Dice coefficient of an anatomical structure, $D_k$, is a metric indicative of plane similarity, given by

$$Dice_k = \frac{2\,|A_Q(k) \cap A_S(k)|}{|A_Q(k)| + |A_S k|} \tag{2}$$

where $k \in \{\text{CP, PVC, CSP, CB}\}$ and the $A_i(k)$ is the number of pixels of structure $k$ present in a plane $i \in \{Q,S\}$. Since the planes cannot be assumed to be parallel, it is expected that there may be a shift in the location of corresponding intracranial structures. To account for this, the intensity images were first rigidly registered and the transformation was applied to the manual segmentation images prior to computing the Dice coefficient.

The mean distance between the surface-extracted and manually-selected planes was computed by extracting the points on the cranial surface which define the contour of the skull in the surface-extracted plane and computing the average Euclidean distance between these points and the manually-extracted plane. The cranial contour points $U_Q$ are defined as the surface points which overlap with the surface-extracted plane, $U_S = (U \cap S)$, where $U \subset N_U$. The Euclidean distance between the points is defined as

$$Dist(Q,S) = \hat{\mathbf{n}}_Q \cdot U_S + p_Q \tag{3}$$

where $\hat{\mathbf{n}} = \dfrac{\mathbf{n}}{|\mathbf{n}|}$ is the unit normal vector, $p_Q$ is the distance of the manually-extracted plane ($Q$) from the origin, and the dot product $\hat{\mathbf{n}}_Q \cdot U_S$ represents the scalar projection of the points on the surface extracted plane ($U_S$) onto plane $Q$.

# 4 Results

The mean and standard deviation of the dihedral angles between the manually-selected and surface-extracted planes are summarized in Table 1, and Figure 3 displays visual examples of surface-extracted plane selection along with their corresponding dihedral angles. It is evident that in all image slices, the relevant anatomical structures were present in the selected planes, regardless of the dihedral angle.

The dihedral angle between $Q$ and $S$ was smallest and least variable for the TT plane: the plane containing the most anatomical landmarks as per the ISUOG guidelines [1]. In contrast, the TC plane displayed the highest variability in plane selection due to the imprecise definition of the angle of 2D image acquisition protocol and anatomical landmarks expected to be visible in the plane (Table 2). The results tabulated in Table 2 exclude 7.69% (4/52 images) of the images which were considered outliers (i.e. $|\theta_d| \geq 25^\circ$ because image does not comply with standard acquisition protocols). These outliers constituted images satisfying any of the following criteria:

- Clinician selected an incorrect plane (due to unclear anatomy visualization caused by motion artifacts)
- Cranial surface was misaligned during the user initialization
- Intracranial structures were indiscernible

The mean Euclidean distances between $Q$ and $S$ planes were all below $5mm$ for all standard planes, which indicates that the surface-extracted planes did not greatly deviate from the planes selected by a clinician for fetal brain assessment.

Table 2 summarizes the segmentation overlap between the manually-selected and surface-extracted planes. Low Dice coefficient values are attributed to the difficulty in achieving proper alignment between non-parallel $Q$ and $S$ planes, undefined manual segmentation protocol for each anatomical structure, and the small size of the intracranial structures on each plane. However, the fact that there is overlap between the structures expected at each plane signifies that the surface-extracted planes contain the appropriate intracranial structures of relevance to clinicians. This highlights that although a surface extracted-plane may not exactly match the clinically-selected plane, it still complies with the ISUOG guidelines [1], meaning that it would be an acceptable choice of plane in a fetal neurosonography evaluation. The results also validate the parametrized surface's ability to extract planes with anatomical consistency in different subjects.

In order to assess clinical usefulness of the surface-extracted planes, we compared head circumference (HC) measurements collected from them (Figure 4a) to those obtained using

the current best clinical method: the mean of three HC measurements collected from different 2D TT US scans on the same fetus (Figure 4b). The difference between the two measurements had a tight fit (i.e. RMSE=18.4$mm$, for HC ranging from 143.4$mm$ to 308.3$mm$ at $18^{+0}$ and $28^{+6}$ weeks, respectively, for our dataset), and the bias error between the two techniques was within the acceptable range of interobserver variability for HC measurements ($\pm$11.1$mm$, 4.8%), i.e. below 5% [5], suggesting clinical viability of the surface-extracted plane measurements.

Table 1: **Plane Similarity Metrics.** Mean and standard deviation values of dihedral angles between manually-selected and surface-extracted planes ($\theta_d(Q,S)$ in degrees), and mean and standard deviation distance values between the planes ($Dist(Q,S)$ in mm).

| Plane | $\theta_d(Q,S)$ | $Dist(Q,S)$ |
|---|---|---|
| TV | $1.5 \pm 8.4°$(n=50) | $4.31 \pm 1.79$ |
| TT | $0.8 \pm 6.6°$(n=51) | $3.55 \pm 1.58$ |
| TC | $2.6 \pm 7.3°$(n=48) | $4.63 \pm 1.91$ |

Table 2: **Plane Overlap Metrics.** Dice Coefficient for anatomical structures of interest ($D_k$). Blank table entries correspond to structures that were consistently absent from the plane.

| Plane | $Dice_{\text{CP}}$ | $Dice_{\text{PVC}}$ | $Dice_{\text{CSP}}$ | $Dice_{\text{CB}}$ |
|---|---|---|---|---|
| TV | $0.509 \pm 0.185$ | $0.334 \pm 0.357$ | $0.195 \pm 0.265$ | — |
| TT | $0.285 \pm 0.314$ | $0.173 \pm 0.349$ | $0.333 \pm 0.387$ | — |
| TC | — | — | $0.136 \pm 0.295$ | $0.097 \pm 0.156$ |

# 5 Conclusion

We have demonstrated the viability of using a topological manifold representation of the fetal skull to extract diagnostic image planes from which clinically-useful information can be extracted. The similarity between the surface-extracted and manually-selected planes relies on the user providing an accurate alignment of surface model and imaged brain. In this paper, we have demonstrated the ability of the parametric surface to extract diagnostic planes containing all of the key landmarks characteristic of a given plane, respecting geometric and anatomical expectations. In addition, the parametrization of the cranial surface allows for anatomical consistency in sampling images from different patients across a wide gestational age window spanning the second and early third trimesters.

Finally, we have shown that biometric measurements obtained from surface-extracted planes are within the accepted range of interobserver variability, implying that the planes would be acceptable in routine assessment of fetal growth. Further experiments include using the 3D nature of the surface model to extract measurements with potential applications to detection of craniosynostosis or craniofacial dysmorphology *in utero* [2].

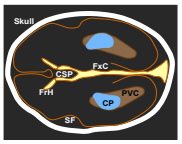
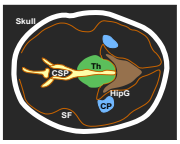
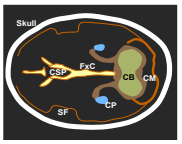
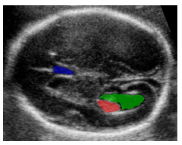
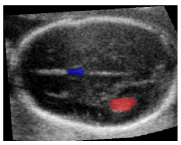
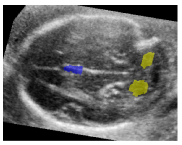
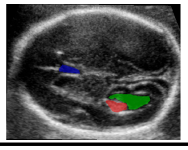
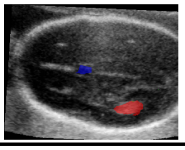
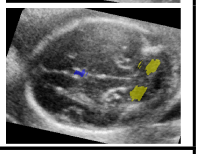
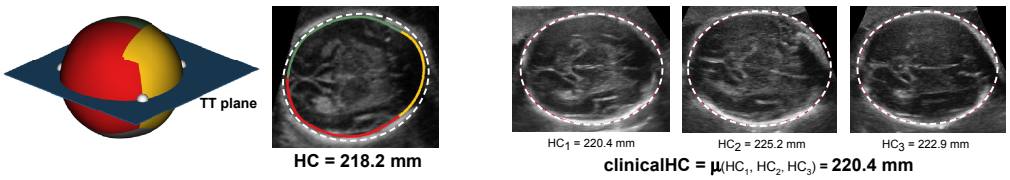
| | TV | TT | TC |
|---|---|---|---|
| Schematic |  |  |  |
| Manually-Selected, $Q$ |  |  |  |
| Surface-Extracted, $S$ |  |  |  |
| $\theta_d(Q,S)$ | 1.1° | -4.1° | 4.6° |

Figure 3: **TV plane selection**. Visual comparison between manually-selected (top row) and surface-extracted TV planes (bottom row) for three different subjects. Manual segmentations of anatomical landmarks are shown: choroid plexus (red), posterior ventricle cavity (green), cavum septum pellucium (blue), and cerebellum (yellow). Dihedral angle between manually-selected and surface-extracted planes is shown below each image pair.




(a) Surface-extracted HC measurement (b) Clinical manual HC measurement

Figure 4: **HC measurement.** Methods of obtaining head circumference measurement from (a) surface-extracted and (b) manual-selected TT planes. For (b), HC was estimated by the mean of clinical annotations from three different images of the same subject.

# References

[1] Sonographic examination of the fetal central nervous system: guidelines for performing the 'basic examination' and the 'fetal neurosonogram'. *Ultrasound Obstet Gynecol*, 29 (1):109–116, 2007.

[2] HC Chen, PY Tsai, HH Huang, HH Shih, YY Wang, CH Chang, and YN Sun. Registration-based segmentation of 3D ultrasound images for quantitative measurement of fetal craniofacial structure. *Ultrasound Med Biol*, 38(5):811–823, 2012.

[3] PE Gill, W Murray, and MH Wright. Practical optimization. 1981.

[4] AIL Namburete, RV Stebbing, and JA Noble. Cranial parametrization of the fetal head for 3D ultrasound image analysis. In *MIUA*, pages 196 – 201, 2013.

[5] I Sarris, C Ioannou, P Chamberlain, E Ohuma, F Roseman, L Hoch, DG Altman, AT Papageorghiou, and INTERGROWTH-21st. Intra- and interobserver variability in fetal ultrasound measurements. *Ultrasound Obstet Gynecol*, 39(3):266–273, 2012.

# Automatic hippocampal multimodal assessment for studies of stroke and small vessel disease

Maria del C. Valdés Hernández[1]
M.Valdes-Hernan@ed.ac.uk
Jaeil Kim[2]
threeyears@kaist.ac.kr
Ian Whitteford[3]
s1104030@sms.ed.ac.uk
Xinyi Qiu[3]
s1108147@sms.ed.ac.uk
Joanna M. Wardlaw[1]
Joanna.Wardlaw@ed.ac.uk
Jinah Park[2]
jinahpark@kaist.ac.kr

[1] Brain Research Imaging Centre
University of Edinburgh
Edinburgh, UK

[2] Department of Computer Science
Korea Advanced Institute of Science and Technology,
Daejeon, South Korea

[3] College of Medicine and Veterinary Medicine
University of Edinburgh
Edinburgh, UK

**Abstract**

We propose a framework for assessing the hippocampi on stroke patients and studies of small vessel disease, where sclerosis, perivascular spaces and infarcts on this structure are common. It includes hippocampal and cavity segmentations, hippocampal shape modelling, feature characterisation and statistical analyses, all which have been particularly developed for assessing extreme abnormalities on this small brain structure on clinical MRI datasets. The hippocampal segmentation uses FSL™ tools. We apply this approach to 48 datasets from a study of mild stroke and assess its relevance on a larger dataset of 189 stroke patients. Using shape similarity metrics we show that the hippocampal shape models generated by our method are accurate on this dataset. We estimate the prevalence and distribution of the features analysed on the samples to discuss the usefulness of our approach.

# 1 Introduction

The temporal lobe and, in particular, the hippocampus play an important role in cognitive processes. Advanced magnetic resonance imaging (MRI) techniques combined with histology have confirmed the presence of markers of small vessel disease in the hippocampus such as microinfarcts and perivascular spaces [1]. Some of them may represent a diffuse vascular process with adverse local effects and/or proxies for larger volumes of infarcts or mild or severe diffuse damage [2]. They appear as cavities of circular or ovoid shape hyperintense in axial T2-weighted [1] and, if >=2-3 mm diameter, hypointense in axial fluid attenuation inversion recovery (FLAIR) MRI sequences. In

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

addition, hippocampal sclerosis occurs in substantial number of older persons. It visually reduces the hippocampal size and causes extreme deformations in its shape.

We developed computational methods for quantitatively assessing hippocampal cavities and hippocampal shape deformities, and integrated them on a framework that gives as output a full characterisation of these structures ready to be used in statistical analyses.

# 2 Methods

## 2.1 Preprocessing

Hippocampi were segmented from T1-weighted volume scans using an atlas-based segmentation approach based on FSL FLIRT and FAST [3], and an ageing template generated in-house from 97 brains of Caucasian individuals aged 65-70 years old. The generated hippocampal binary masks were all visually inspected and manually rectified using Analyze 12.0$^{TM}$. FLAIR hypointensities in the hippocampi, of >=2mm diameter with circular or oval shapes, were, separately, semi-automatically masked also using the same software. They correspond to hyperintensities in T2-weighted and have been reported as small cavities [1]. Results were overseen by a trained image analyst and a neuroradiologist.

## 2.2 Hippocampal shape modelling

We developed a shape modelling method tailored to small and highly-variable shape structures like the hippocampus, based on the Laplacian surface deformation framework, proposed by Kim and Park [4]. This method is characterised by a non-rigid template deformation in a coarse-to-fine style to minimise the distortion of the point distribution of the template surface model during the target surface reconstruction. It encodes the point distribution of the template surface model as Laplacian coordinates representing the relative positions of each point with respect to the average position of their neighbours and preserves the Laplacian coordinates of the template model as rigid as possible to sample the target surface with the point distribution of the template model. Consequently, the preservation of this point-wise correspondence allows us to make comparative analyses on shape variations.

From the binary hippocampal masks of the sample, mean shape images are constructed via an iterative and non-linear image registration between the binary masks following the process described in [5]. From the mean shape images, the template surface models for left and right hippocampi are, then, generated as triangular meshes using spherical harmonics following by a point distribution model (SPHARM-PDM v1.11) [6]. Further, to recover each individual shape, the template models are rigidly aligned to each target volume (i.e. left and right hippocampal binary masks) via the iterative closest point algorithm [7], following non-rigid deformation. Under the definition of Laplacian coordinates as the difference between each point's position and the centre of mass for their neighbouring points, the surface $\overline{\mathbf{V}}$ of the non-rigid template deformation is formulated as a quadratic form (equation (1)).

$$\overline{\mathbf{V}} = \underset{\mathbf{V}'}{\operatorname{argmin}} \left( \left\| \alpha \left( \mathbf{LV}' - \mathbf{LV} \right) \right\|^2 + \sum_{1}^{n} \left\| b_i - v_i \right\|^2 \right) \quad (1)$$

where $\mathbf{V}$ is the set of all vertices $v_i$ for the points (each point denoted as $i$), α is a weighting term that controls the rigidity of the deformable model, $L$ is a discrete Laplacian operator and $b_i$ is the closest image boundary on the surface normal at each point, defined as:

$$b_i = \beta \cdot \left( m_i - v_i \right) + v_i \quad (2)$$

where $m_i$ is the closest image boundary on the direction normal to the vertex and $\beta$ is a weighting term introduced to regularise the vertex transformation.

In order to preserve the Laplacian coordinates as rigid as possible, we employ a progressive weighting scheme for α as described in [4]. Under this scheme, the template surface model is deformed with as-large-as-possible α values until the points of the template model best fit the boundary. This iterative process decreases α in a stepwise way, together with the magnitude of the displacement of each vertex. We experimentally determined that the maximum α value for our purpose was equivalent to 10 times the volumetric ratio between template and target. The optimal $\mathbf{V}$ for equation (1) can be obtained by solving a linear system using a matrix form of the Laplacian operator via a linear least squares approach. The solution is described in detail elsewhere [4].

We added the rotation and scale invariant (RSI) transformation proposed by [8] to the modelling process described in [4]. It constrains the vertex transformations only to rotation, isotropic scale and translation. This helps regularising the individual vertex transformation, derived by external factors, to the transformations of the neighbour vertices using them as reference.

### 2.3 Analyses of shape and cavities

A shape deformity map from the individual surface models is generated using the mean surface models of left and right hippocampi. The individual surface models are normalised via an isotropic rescaling of each shape model using hippocampal size and the generalised Procrustes analysis [9]. Local shape differences are determined by the displacement vectors between the corresponding vertices of the individual surface models and the mean surface model. The shape deformity at each vertex is computed as the signed Euclidean norm of the displacement vectors, projected on the vertex normal on the mean surface model to determine the direction of local shape changes.

The number of hippocampal cavities (i.e. hippocavities) and their load per hippocampi are derived automatically by assessing the intersection between the hippocavities mask and the hippocampal binary masks after a "fill-hole" operation is performed. The count was determined by counting the isolated clusters of hippocavities on each hippocampal mask. The load per structure is the total relative hippocavity volume of all hippocavities within a hippocampus. The morphology of individual hippocavities is quantified by its volume, maximum in-plane extent, roundness and sphericity. The "clusterisation" of small individual hippocavities is determined through the assessment of "compactness" as defined by Bibriesca in [10].

## 3 Experiments and Results

## 3.1 Datasets

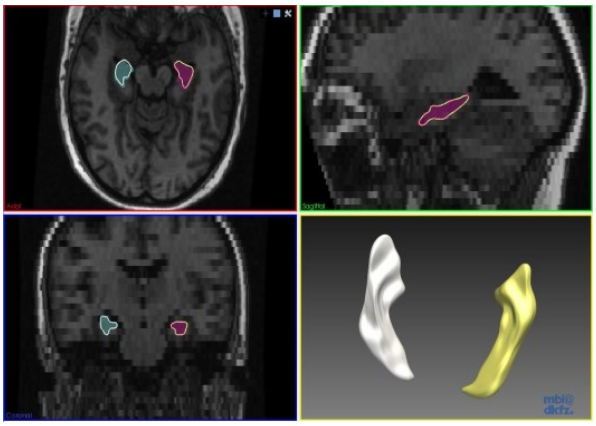
We used MRI data from patients with lacunar or mild cortical stroke that presented with stroke symptoms to a teaching hospital between 2002 and 2013 and formally consented to participate on studies of stroke. All MRI data were acquired on the same 1.5T GE Signa Horizon HDxt clinical scanner operating in research mode with a self-shielding gradient set with maximum gradient of 33 mT/m, and an 8-channel phased-array head coil.

From a sample of 48 patients (14 women) mean age 66 (SD=10) years, we assessed the hippocampi and cavities as explained previously. To generate hippocampi binary masks we used T1-weighted spin-echo MRI scans of which 6 had matrix dimensions of 256x216x256 and voxel size of 1.0156x0.9x1.0156 $mm^3$ and the rest had matrix dimensions of 256x256x42 and voxel size of 0.9375x0.9375x4 $mm^3$. Cavities were assessed using FLAIR and T2-weighted images, only on 42/48 subsets in which images had matrix dimensions of 256x256x42 and voxel size of 0.9375x0.9375x4 $mm^3$. In the rest (i.e. 6/48 datasets), these sequences had very anisotropic voxels (0.4688x0.4688x6 $mm^3$) that were unsuitable for this purpose.

We visually assessed the presence of hippocavities on T2-weighted and FLAIR MR images from 189 MRI datasets from patients with acute lacunar ischaemic stroke clinical syndrome and a relevant lesion visible on the diffusion-weighted image, to examine their prevalence, appearance and distribution.

## 3.2 Results of the analyses of hippocampal shape and cavities on a subset of 50 stroke patients

In the shape modelling, β was set at 0.3 and α was progressively reduced from 30 to 1. For the 6 datasets with nearly isotropic voxels, β=0.5 and minimum α=0.3 were preferred, but for cross-evaluation purposes the results obtained with these α and β values were discarded. We visually checked the results using the MITK Workbench 2013.06.00 (Figure 1(a)) and calculated three similarity measures to quantitatively evaluate the modelling results (Table 1).



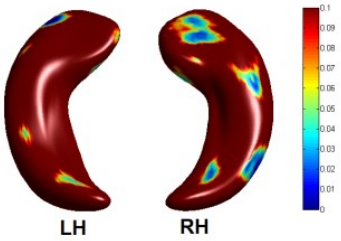



Figure 1: Illustration of results obtained from the shape analysis: Screenshot of the hippocampi shape models (superimposed onto the binary masks and the T1-weighted image) for a subject (left), and regions of significant deformation differences between the hippocampi from patients that had cortical vs. lacunar stroke (pattern superimposed onto the mean shapes of left and right hippocampi) (right).

| Measure | Right Hippocampus | Left Hippocampus |
|---|---|---|
| Dice coefficient | 0.9135 (IQR 0.0338) | 0.8867 (IQR 0.0342) |
| Hausdorff distance (mm) | 4.6482 (IQR 3.3304) | 7.8551 (IQR 2.5206) |
| Mean distance (mm) | 0.2006 (IQR 0.1711) | 0.3595 (IQR 0.2450) |

Table 1: Median and interquartile range values of the shape similarity measurements between the surfaces of the individualised (i.e. fitted) shape models and the hippocampal binary masks (after converting the latter to voxel meshes).

FLAIR hypointensities with intensity level 0 were found in 9/42 datasets, of which in 4/9 were single voxels. In general, single voxel FLAIR hypointensities, were found in 14/42 datasets, either scattered within the structure or in its periphery, or distributed in a line along the Cornu Ammonis 1 (dorsal hippocampi). Multivoxel structures of a single intensity value (of which 21 were elongated and 24 were ovoid) appeared on 25/42 datasets. Multivoxel structures of different intensity and shape patterns appeared on 37/42 datasets (Figure 2). The maximum hippocavity volume found was 49.22 $mm^3$ (ovoid shape).

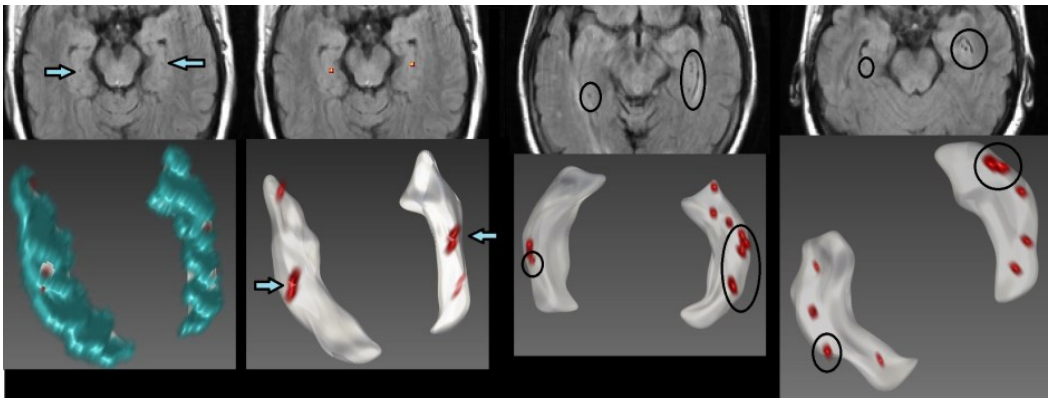


Figure 2: Three examples of the distribution of FLAIR hypointensities/T2-weighted hyperintensities on representations of their respective hippocampi. Arrows point to hippocavities penetrating the hippocampi. Encircled are hippocavities on the surface. The blue surface is the 3D representation of the binary mask.

### 3.3 Incidence of hippocampal cavities on a large sample of patients with lacunar stroke

Hippocavities were observed on 86/99 patients with the index stroke lesion in the left (L) cerebral hemisphere, on 67/78 patients with the index stroke in the right (R) cerebral hemisphere and, in general, on 82% of the sample (155/189), with a median value of 2 cavities per hemisphere regardless of the location of the index stroke. Only on two patients a hippocavity ovoid in shape with maximum diameter of 3-4 mm was observed. The rest were either round with 1-2 mm diameter or elongated small structures coincident with the appearance of perivascular spaces [11].

## 4 Discussion and Conclusion

An integrated framework that includes shape modelling and cavity characterisation for the assessment of the hippocampus on clinical MRI datasets of stroke patients and patients with small vessel disease has been proposed here. The anisotropy of the imaging voxels, extreme structural deformations and prevalence of several abnormalities imposed a manual

editing step for obtaining hippocampal binary masks and cavities. The rest of the process was fully automatic. The shape modelling pipeline presented, preserved the individual shape details allowing the detection of morphological changes of this structure. Further work will involve 1) evaluating its robustness against other state-of-the-art shape modelling methods and on detecting subtle morphological changes of this structure and 2) combining the hippocampal and cavity masks using the latter as a constraint on the deformation process.

Not all FLAIR hypointensities computationally identified by thresholding had the characteristics of those visually identified as hippocavities on the larger clinical sample. However, with the proposed framework, these could be identified. The prevalence and characteristics of those identified as possible hippocavities on the sample computationally analysed were in agreement with the visual assessment on the larger sample and with existent clinical reports [1,2].

# Acknowledgements

This work was funded by Row Fogo Charitable Trust, the National Research Foundation of Korea (Grant no. 2012K2A1A2033133/no.2011-0009761) and the College of Medicine and Veterinary Medicine at the University of Edinburgh, UK.

# References

[1] S.J. van Veluw, L.E.M. Wisse, H.J. Kuijf *et al*. Hippocampal T2 hyperintensities on 7 Tesla MRI. *NeuroImage:Clinical*, 3:196-201, 2013.

[2] Z. Arvanitakis, S.E. Leurgans, L.L. Barnes *et al*. Microinfarct pathology, dementia, and cognitive systems. *Stroke*, 42 (3):722-727, 2011.

[3] B. Patenaude, S. Smith et al. FMRIB Technical Report TR07BP1. Oxford, 2007.

[4] J. Kim and J. Park. Organ shape modelling based on the laplacian deformation framework for surface-based morphometry studies. *J Comp Sci Eng*, 6:219-226, 2012.

[5] G. Heitz, T.Rohlfing, C.R. Maurer Jr. Statistical shape model generation using non rigid deformation of a template mesh. *Sig Proc* 1411-1421, 2005.

[6] M. Styner, I. Oguz, S Xu, et al. Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM. *Insight Journal. NIH Public Access*, (1071):242-250, 2006.

[7] P.J. Besl and H.D. McKay. A method for registration of 3D shapes. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 14(2):239-256, 1992.

[8] O. Sorkine, D. Cohen-Or, Y Lipman et al. Laplacian surface editing. Proc. Of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, 175-184, 2004.

[9] J.C. Gower. Generalized Procrustes Analysis. *Psychometrika*, 40:33-51,1975.

[10] E. Bribiesca. An easy measure of compactness for 2D and 3D shapes. *Pattern Recognition*, 41:543-554, 2008.

[11] J.M. Wardlaw, E.E. Smith, G.J. Biessels et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol, 12:822-838, 2013.

# Fast Template-based Shape Analysis using Diffeomorphic Iterative Centroid

Claire Cury[(a,b,c,d,e)1]
claire.cury.pro@gmail.com
Joan Alexis Glaunès[2]
alexis.glaunes@mi.parisdescartes.fr
Marie Chupin[(a,b,c,d,e)1]
marie.chupin@upmc.fr
Olivier Colliot[(a,b,c,d,e)1]
olivier.colliot@upmc.fr
The Alzheimer's Disease Neuroimaging Initiative[0]
.

[1] *a*- Inserm U 1127,
F-75013, Paris, France
*b*- CNRS UMR 7225,
F-75013, Paris, France
*c*- Sorbonne Universités, UPMC Univ Paris 06, UMR S 1127,
F-75013, Paris, France
*d*- Institut du Cerveau et de la Moelle épiniére, ICM,
F-75013, Paris, France
*e*- Inria Paris-Rocquencourt, Aramis project-team,
F-75013, Paris, France

[2] MAP5, Université Paris Descartes, Sorbonne Paris Cité, France

**Abstract**

A common approach for the analysis of anatomical variability relies on the estimation of a representative template of the population, followed by the study of this population based on the parameters of the deformations going from the template to the population. The Large Deformation Diffeomorphic Metric Mapping framework is widely used for shape analysis of anatomical structures, but computing a template with such framework is computationally expensive. In this paper we propose a fast approach for template-based analysis of anatomical variability. The template is estimated using a recently proposed iterative approach which quickly provides a centroid of the population. Statistical analysis is then performed using Kernel-PCA on the initial momenta that define the deformations between the centroid and each subject of the population. This approach is applied to the analysis of hippocampal shape on 80 patients with Alzheimer's Disease and 138 controls from the ADNI database.

## 1 Introduction

Computational Anatomy aims at developing tools for the quantitative analysis of variability of anatomical structures, and its variation in healthy and pathological cases [8]. A common approach in Computational Anatomy is template-based analysis, where the idea is to compare anatomical objects variations relatively to a common template. These variations are analysed using the ambient space deformations that match each individual structure to

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

the template. The Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework [1] has been widely used for the study of the geometric variation of human anatomy. This framework generates metrics between deformable shapes and provides smooth and non ambiguous matchings between objects.

Estimating a template from the population in the LDDMM framework, to be used for further template-based statistical analysis, is a computationally expensive task. Several methods have been proposed. Vaillant et al. proposed in 2004 [12] a method based on geodesic shooting which iteratively updates a shape by shooting towards the mean directions. The method proposed by Glaunès and Joshi [7] starts from the whole population and estimates a template by co-registering all subjects using a backward scheme. A different approach was proposed by Durrleman et al. [4, 5]. The method initializes the template with a standard shape, and uses a forward scheme: deformations are defined from the template to the subjects. The method presented by Ma et al. [9] uses an hyper template which is an extra fixed shape, and optimizes at the same time deformations from the hyper template to the template and deformations from the template to subjects of the population. All these methods are expensive in terms of computation time, due to the iterations needed for the convergence of the method.

In this paper, we propose a fast approach for template-based shape analysis in the LDDMM framework. The template is estimated using an iterative approach which quickly provides a centroid of the population [3]. This method iteratively computes a centroid of the population in the LDDMM framework, which requires only $N-1$ matching steps, with $N$ the number of subjects, while template estimation methods typically require $N$ such steps per iteration. Using this centroid computation as initialization drastically reduces the computation time for template computation. We use this iterative centroid as a template to study the anatomical variability by analysing the deformations from the centroid to the subjects of the population using Kernel-PCA [10]. This approach was applied to 218 hippocampi (80 patients with Alzheimer's Disease, 138 controls) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database [1].

# 2 Methods

## 2.1 LDDMM framework

We briefly present the LDDMM framework to introduce notations, for more details see [1]. Deformation maps $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ are generated via integration of time-dependent vector fields $v(x,t), x \in \mathbb{R}^3, t \in [0,1]$, such that each $v(\cdot,t)$ belongs to a Reproducing Kernel Hilbert Space $V$ with kernel $K_V$. The transport equation

$$\begin{cases} \frac{d\phi_v}{dt}(x,t) = v(\phi_v(x,t),t) & \forall t \in [0,1] \\ \phi_v(x,0) = x & \forall x \in \mathbb{R}^3 \end{cases} \tag{1}$$

has a unique solution, and one sets $\varphi_v = \phi_v(\cdot,1)$ the diffeomorphism induced by $v(x,t)$. In a discrete setting, optimal vector fields $v(x,t)$ are expressed as combinations of spline fields: $v(x,t) = \sum_{p=1}^{n} K_V(x,x_p(t))\alpha_p(t)$, where $x_p(t) = \phi_v(x_p,t)$ are the trajectories of control points $x_p$ (the vertices of the mesh to be deformed), and $\alpha_p(t) \in \mathbb{R}^3$ are time-dependent vectors called momentum vectors. Optimal trajectories between shapes can be shown to

[1] (adni.loni.usc.edu). The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California - San Francisco.

satisfy geodesic equations for a metric on the set of control points [12]. As a result the full deformation between a template shape and its surface target $S_i$ is encoded by the vector of initial momentum vectors $\alpha^i(0) = (\alpha_p^i(0))_{1\leq p\leq n_i}$ located on the vertices of the template mesh.

## 2.2 Iterative Centroid method

We proposed in 2013 [3] an Iterative Centroid method which use the LDDMM framework and the framework of currents [11]. The authors propose three different schemes for centroid computation, but here we only use the first one, which is the faster one. The Iterated Centroid method consists in applying the following procedure: given a collection of N shapes Si, we first set $C1 = S1$ (initial centroid) and successively update the centroid by matching it to the next shape and moving it along the geodesic flow. (Algorithm 1).

**Data**: $N$ surfaces $S_i$
**Result**: 1 surface $C_N$ representing the centroid of the population
$C_1 = S_1$;
**for** *i from* 1 *to* $N-1$ **do**
  $C_i$ is matched to $S_{i+1}$ which results in a deformation map $\phi_{v^i}(x,t)$;
  Set $C_{i+1} = \phi_{v^i}(C_i, \frac{1}{i+1})$ which means we transport $C_i$ along the geodesic and stop at time $t = \frac{1}{i+1}$;
**end**

**Algorithm 1**: Iterative Centroid algorithm

This algorithm can be defined in a pure Riemannian setting, for the averaging of a set of points $p_k$ (instead of shapes $S_i$) on a Riemannian manifold $M$. If points $p_k$ belong to a vector space and the metric is flat, the algorithm converges towards the mean $\frac{1}{N}\sum_k p_k$, but in general it depends on the ordering of the points. Back to our shape space framework, we showed in a previous study [3] that indeed the ordering changes the result, but the different results are very close to each other. Emery and Mokobodzki [6] proposed to define the centroid not as a unique point but as the set $\mathscr{C}_N$ of points $p \in M$ satisfying $f(p) \leq \frac{1}{N}\sum_{k=1}^{N} f(p_k)$, for any convex function $f$ on $M$ (a convex function $f$ on $M$ being defined by the property that its restriction to all geodesics is convex). This set $\mathscr{C}_N$ takes into account all centroids obtained by bringing together points $p_k$ by all possible means, i.e. recursively by pairs, or by iteratively adding a new point, as we are doing with Algorithm 1. To further reduce the computational load, we used a GPU implementation for the kernel convolutions involved in the matchings.

## 2.3 Statistical Analysis

For the statistical analysis of shapes, we compute a centroid of the population, and we use it as a template. Then we deform the centroid toward each shape of the population to obtain the deformations from the centroid to the population. The deformations are determined by the vector of initial momentum vectors $\alpha^i(0)$; we can analyse these deformations with a Kernel-PCA [10] as in [12]. The Kernel-PCA is the non-linear version of the Principal Component Analysis (PCA) which can be used for learning shape variability from a training set, or for reducing dimensionality of the shape space. In fact the difference between the two versions is in the computation of the covariance matrix, which writes

$Cov(i,j) = \frac{1}{N-1}(\alpha^i(0) - \bar{\alpha}(0))K_V(x)(\alpha^j(0) - \bar{\alpha}(0))$, with $\bar{\alpha}(0)$ the mean of initial momentum vectors and $K_V(x)$ the matrix of the $K_V(x^i, x^j)$. Then we can run a standard PCA on the covariance matrix $Cov$.

# 3 Experiments and results

## Dataset

The method was applied to the analysis of hippocampal shape of 80 patients with Alzheimer's Disease (AD) and 138 controls (CN) ($N = 218$) from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Left hippocampi were segmented with the SACHA software (Chupin et al. [2]) from 3D T1 weighted MRI. Then the meshes were computed using the BrainVisa [2] software. They are composed of 800 vertices on average.

## Results

We computed a centroid for each of the two populations (AD and CN) using Algorithm 1. The two centroids are denoted $IC(AD)$ and $IC(CN)$. Computation times were 2.4 hours for $IC(AD)$ and 3.6 hours for $IC(CN)$. To assess whether the centroids are close to the center of the respective population, we computed the ratio $R = \frac{\| \frac{1}{N} \sum_{i=1}^{N} v_0(S_i) \|_V}{\frac{1}{N} \sum_{i=1}^{N} \| v_0(S_i) \|_V}$ between the mean of the norms of initial vector fields from the centroid to the population and the norm of the mean of initial vector fields. Both ratios are 0.25, which means that both centroids are correctly centred even though they are not exactly at the Fréchet mean (which would correspond to $R = 0$). To visualise differences between $IC(CN)$ and $IC(AD)$, we computed distances between vertices (figure 1) of $IC(CN)$ and the deformation of $IC(CN)$ on $IC(AD)$.

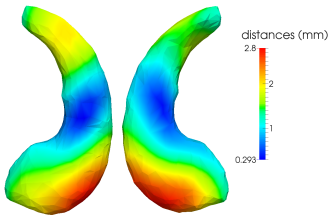



Figure 1: Distances between $IC(CN)$ and $IC(AD)$. On the left, the hippocampus is viewed from below.

We then analysed the variability of the AD and the CN populations using Kernel-PCA. Figure 2 shows, for each group, the principal mode of variation. This figure is obtained by geodesic shooting from each centroid in the first principal direction with a magnitude of $\pm 2\sigma$. One can note that, while the templates of the two groups are different, the variabilities of both groups share similarities. Nevertheless, there seems to be less variability in the medial part of the body for the CN group.

In order to visualize the localization of $IC(CN)$ and $IC(AD)$ and their modes within the whole population, we computed a centroid of the whole population and performed the corresponding K-PCA. We then projected $IC(CN)$ and $IC(AD)$ on the two first principal

[2] http://www.brainvisa.info

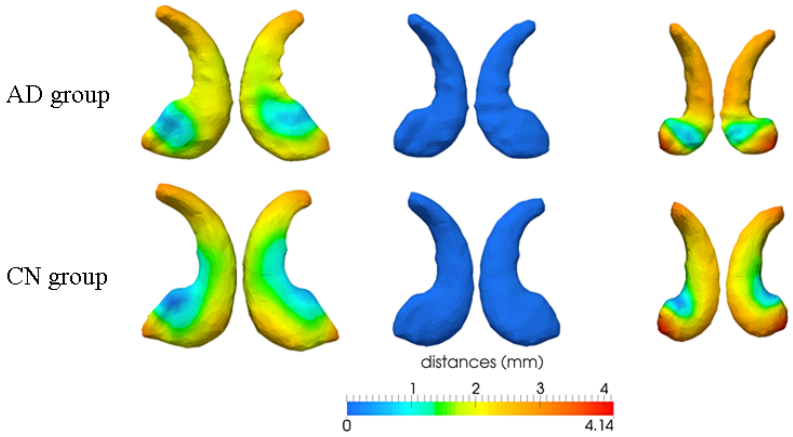



Figure 2: First mode of deformation of the AD group (top) and of the CN group (bottom). For each row, the centroid is in the center (in blue), on the right its deformation at $+2\sigma$, and at $-2\sigma$ on the left. The colormap indicates the displacement of each vertex between the corresponding centroid and its deformation.

components. We can observe (figure 3) that $IC(CN)$ is on the left of the global centroid, and $IC(AD)$ is on the right, and the 3 principal modes of variation have different directions.

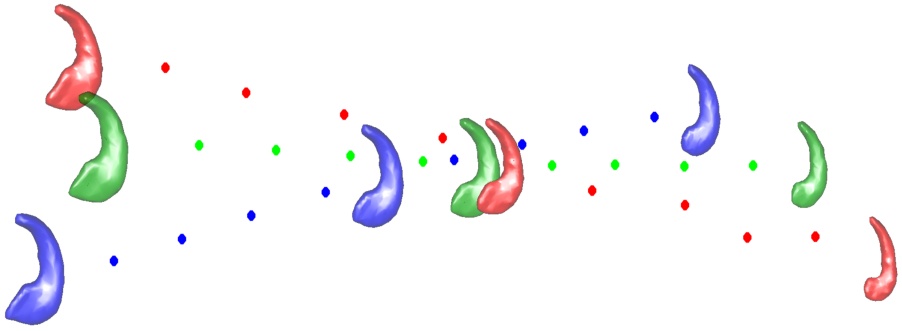


Figure 3: First principal modes of the whole population (in green). In blue, the projections of $IC(CN)$ and projections of its deformations at $\pm 2\sigma_{CN}$ in the direction of its first mode of variation. In red, for the AD group. Intermediate points are show the trajectory of the first modes.

# 4 Conclusion

In this paper, we proposed a new approach for fast template-based shape analysis in the LDDMM framework. The template is estimated using a diffeomorphic iterative centroid method. The Iterative Centroid is roughly centred within the population of shapes. Analysis of variability is then based on a Kernel-PCA of the initial momenta that define the deformations of the template to each individual. We applied this approach to the analysis of hippocampal shape in AD patients and control subjects. Projection of the templates of the CN group and the AD group onto the main modes of variability of the whole population show that they are located on the main axis of variability. The template estimation takes only a

couple of hours for both populations (with about 100 subjects each). The low computational load of the proposed approach makes it applicable to very large datasets.

### Acknowledgments

The research leading to these results has received funding from ANR (project HM-TC, grant number ANR-09-EMER-006, and project KaraMetria, grant number ANR-09-BLAN-0332) and from the program Investissements d´ avenir ANR-10-IAIHU-06.

## References

[1] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005.

[2] M. Chupin, A. Hammers, R. S. N. Liu, O. Colliot, J. Burdett, E. Bardinet, J. S. Duncan, L. Garnero, and L. Lemieux. Automatic segmentation of the hippocampus and the amygdala driven by hybrid constraints: Method and validation. *NeuroImage*, 46(3): 749–761, 2009.

[3] C. Cury, J. A. Glaunés, and O. Colliot. Template Estimation for Large Database: A Diffeomorphic Iterative Centroid Method Using Currents. In Frank Nielsen and Frédéric Barbaresco, editors, *GSI*, volume 8085 of *LNCS*, pages 103–111. Springer, 2013.

[4] S. Durrleman, X. Pennec, A. Trouvé, N. Ayache, et al. A forward model to build unbiased atlases from curves and surfaces. In *2nd Medical Image Computing and Computer Assisted Intervention. Workshop on M.F.C.A.*, pages 68–79, 2008.

[5] S. Durrleman, M. Prastawa, J. R. Korenberg, S. Joshi, A. Trouvé, and G. Gerig. Topology Preserving Atlas Construction from Shape Data without Correspondence Using Sparse Parameters. In *MICCAI 2012*, LNCS, pages 223–230. Springer, 2012.

[6] M. Emery and G. Mokobodzki. Sur le barycentre d'une probabilité dans une variété. In *Séminaire de probabilités*, volume 25, pages 220–233. Springer, 1991.

[7] J. A. Glaunès and S. Joshi. Template estimation from unlabeled point set data and surfaces for Computational Anatomy. In *Proc. of the International Workshop on the MFCA-2006*, pages 29–39, 2006.

[8] U. Grenander and M. I. Miller. Computational anatomy: An emerging discipline. *Quarterly of applied mathematics*, 56(4):617–694, 1998.

[9] J. Ma, M. I. Miller, A. Trouvé, and L. Younes. Bayesian template estimation in computational anatomy. *NeuroImage*, 42(1):252–261, 2008.

[10] B. Schölkopf, A. Smola, and K.-R. Müller. Kernel principal component analysis. In *Artificial Neural Networks—ICANN'97*, pages 583–588. Springer, 1997.

[11] M. Vaillant and J. A. Glaunés. Surface matching via currents. In *Information Processing in Medical Imaging*, pages 381–392. Springer, 2005.

[12] M. Vaillant, M. I. Miller, L. Younes, and A. Trouvé. Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23:S161–S169, 2004.

# Analysis of Cellular Images

# Trainable segmentation of phase contrast microscopy images based on local Basic Image Features histograms

Nicolas Jaccard[1]
n.jaccard@ucl.ac.uk
Nicolas Szita[2]
n.szita@ucl.ac.uk
Lewis D. Griffin[1]
l.griffin@cs.ucl.ac.uk

[1] Department of Computer Science / CoMPLEX
University College London

[2] Department of Biochemical Engineering
University College London

**Abstract**

Phase contrast microscopy (PCM) is routinely used for the inspection of adherent cell cultures in all fields of biology and biomedicine. Key decisions for experimental protocols are often taken by an operator based on typically qualitative observations. However, automated processing and analysis of PCM images remain challenging due to the low contrast between foreground objects (cells) and background as well as various imaging artefacts. We propose a trainable pixel-wise segmentation approach whereby image structures and symmetries are encoded in the form of multi-scale Basic Image Features local histograms and classification of them is learned by random decision trees. This approach was validated for segmentation of cell versus background, and discrimination between two different cell types. Performance close to that of state-of-the-art specialised algorithms was achieved despite the general nature of the method. The low processing time (<4s per $1280 \times 960$ pixel images) is suitable for batch processing of experimental data as well as for interactive segmentation applications.

## 1 Introduction

Phase contrast microscopy (PCM) is widely used as the *de facto* light microscopy modality for the inspection of adherent cell cultures. Segmentation of PCM images is challenging due to the low contrast between the cell objects and the image background: generic thresholding approaches (e.g. Otsu's [9]) do not usually produce satisfactory results. Specialised approaches for PCM image segmentation that rely on *a priori* knowledge of the structure and properties of the images have been developed, including methods based on contrast filters [2, 6, 7, 14], active contours [1, 12], weak watershed assemblies [3], and image formation models [16]. More recently, trainable segmentation methods for microscopy images based on statistical learning of image features (e.g. intensity, texture) have been gaining traction [8, 13, 15]. Random forest classifiers were found to be suitable for segmentation tasks due to low computational complexity and their ability to accommodate large datasets such as images [11, 13]. Typically, trainable segmentation involves using the responses to a bank

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

of linear and non-linear filters computed at multiple scales as feature vectors for pixel-wise classification. However, the vector for a given pixel typically contains only a single value per scale for a given feature and thus does not encode for potentially valuable local information and context.

In this contribution, we describe a framework for PCM image segmentation whereby local histograms encoding image features at multiple scales were used as the input to random decision trees classifiers. This was achieved by computing Basic Image Features (BIFs), an image representation whose pixels take one of seven values depending on local features and symmetries [4]. This small range of possible pixel values allowed efficient construction of local histograms and classifier training was computationally tractable even in the case where multiple scales were considered. The segmentation performance is assessed using two separate PCM images dataset with distinct challenges. It is also compared to specialised PCM segmentation algorithms.

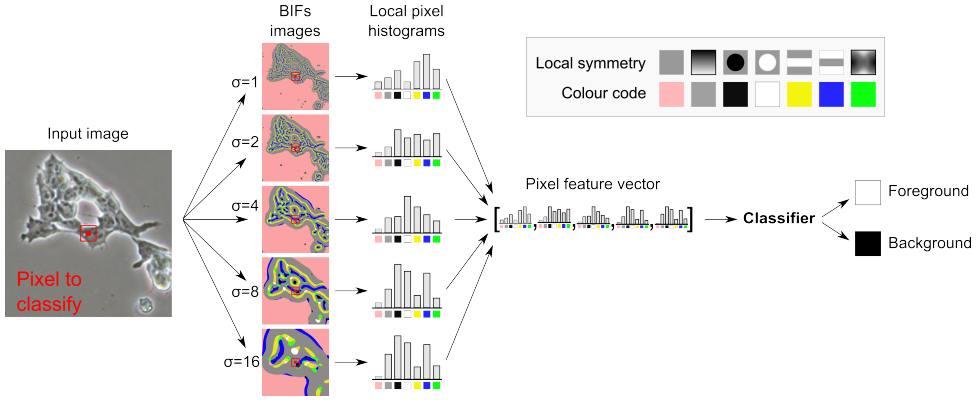



Figure 1: PCM pixel classification based on local histograms of BIFs.

# 2 Trainable segmentation

## 2.1 General approach

PCM images were segmented based on local histograms of Basic Image Features (BIFs) (Fig. 1). First, BIFs of the input image were computed at various scales. Local BIFs histograms were then computed for windows centred at each pixel of the image. The feature vector for classification was constructed by concatenation of the local BIFs histograms obtained for a given pixel of the input across all scales considered. The dimensions of the pixel feature vectors were thus $M \times 7$ where $M$ is the number of scales considered. For comparison purposes, the situation where a single value per scale per pixel was considered, effectively corresponding to a window width of 0 pixel. Pixel feature vectors for classification were then of dimensions $M \times 1$.

The classifier used was a random forest with 20 trees and $\sqrt{F}$ features sampled at each split where $F$ is the total number of features. The output of the classifier was a binary label, with 1 for foreground objects (i.e. cells) and 0 for image background. This output was used as is for segmentation without further processing or refinement.

## 2.2 Local Basic Image Features histograms computation

The computation of Basic Image Features (BIFs) consisted in classifying the output obtained from convolution of an image with a bank of derivative-of-Gaussian (DtG) filters into one of seven categories. These categories corresponded to distinct local image structures, as defined by local symmetries [4]. The response of the convolution of the image $I$ with one of the DtG filter was denoted $c_{ij}$ where $i$ and $j$ represented the order in the x and y directions, respectively. Scale normalised response $s_{ij}$ was then computed as shown in equation 1.

$$s_{ij} = \sigma_B^{i+j} c_{ij} \tag{1}$$

Based on the scale normalised response, an intermediate calculation is carried out as shown in equations 2 and 3.

$$\lambda = s_{20} + s_{02} \tag{2}$$

$$\gamma = \sqrt{(s_{20} + s_{02})^2 + 4s_{11}^2} \tag{3}$$

Both $\lambda$ and $\gamma$ were computed for each pixel of the input image $I$. Pixels were then classified in one of seven categories based on the largest of $\{\varepsilon c_{00}, \sqrt{c_{10}^2 + c_{01}^2}, \lambda, -\lambda, \frac{\gamma+\lambda}{\sqrt{2}}, \frac{\gamma-\lambda}{\sqrt{2}}, \gamma\}$, resulting in a BIFs image $I_B$. BIFs computation was thus controlled by two parameters: the scale (standard deviation) $\sigma_B$ of the DtG filters and a value $\varepsilon$ that controls when a pixel should be considered flat (i.e. no specific structure). For this work, $\varepsilon$ was kept at a constant value of 0.03, which was empirically found to produce good results regardless of the feature scale considered.

Soft-edged local BIFs histograms were computed by convolution [5]. First, seven binary masks $b^{(k)}$ were generated as shown in equation 4, one per Basic Image Feature.

$$b^{(k)}(x,y) = \begin{cases} 1 & \text{if } I_B(x,y) = k \\ 0 & \text{otherwise} \end{cases} \quad \text{for },k = 1,2,\ldots,7 \tag{4}$$

Images $C^{(k)}$ were obtained by convolution of each binary mask $b^{(k)}$ with a Gaussian kernel $G_{\sigma_w}$ of standard deviation $\sigma_w$ equal to the desired window size as shown in equation 5.

$$C^{(k)}(x,y) = G_{\sigma_w} * b^{(k)} \text{ for } k = 1,2,\ldots,7 \tag{5}$$

The histogram at location $(x,y)$ was then constructed by concatenating the values of obtained across the seven $C^{(k)}$ images for that location, as shown in equation 6.

$$H(x,y) = [C^{(1)}(x,y), C^{(2)}(x,y), \ldots, C^{(7)}(x,y)] \tag{6}$$

## 2.3 Datasets and segmentation performance evaluation

Two datasets were used for segmentation performance evaluation (Fig.2). The first one was a set of 50 250×250 pixel mouse embryonic stem cells (mESC) PCM images [6]. This dataset was used to evaluate the performance of the algorithm for a simple foreground versus background segmentation task. The second dataset comprised 20 500×500 pixel PCM images of human embryonic stem cells (hESC) co-cultured with mouse embryonic fibroblasts (MEFs). This dataset was used to evaluate algorithm performance for the discrimination between two

foreground object types with similar visual features. This second dataset was used in a previous study [10] where a preliminary, unoptimised implementation of the approach described here-in resulted in promising results but at the cost of long processing times (∼40s per images). Due to the nature of the cells imaged, it was not possible to segment individual cells. Instead, the goal was the classification of pixels as either foreground or background.

Segmentation performance was evaluated by comparison of the algorithm output with ground truth images annotated by human experts. The agreement between the two was calculated using the F-score (i.e. Dice's coefficient). A leave-one-out cross-validation (LOOCV) approach was taken whereby the classifier was trained using 50000 pixels randomly sampled across $N-1$ images before being used to predict the labels for each pixel of the left out image. This was repeated $N$ times so that all images were left out once. The reported LOOCV F-score was thus the mean F-score across the $N$ images.

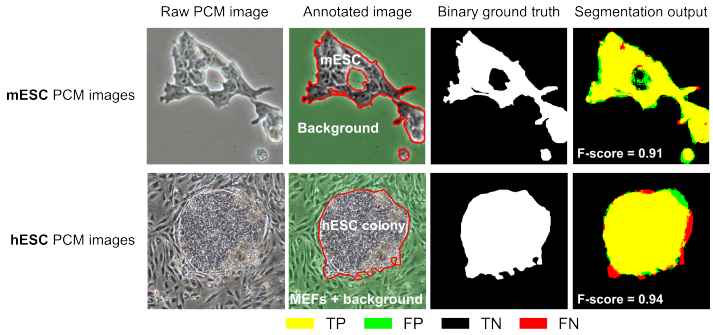



Figure 2: Datasets used for segmentation performance evaluation: mouse embryonic stem cells (mESC) and human embryonic stem cells (hESC) PCM images. The last column shows the agreement between the segmentation output using optimal parameters and the ground truth.

# 3 Segmentation performance

Segmentation performance was evaluated for both the mESC and hESC PCM images datasets over a range of parameter values. The width of the local histogram window ($\sigma_w$) was varied between 5 and 400 pixels. Five base BIFs scales ($\sigma_B$) were considered: 1, 2, 4, 8, and 16. All 31 permutations of these scales were evaluated for each window size $\sigma_w$. For both datasets, the best performance was obtained using local BIFs histograms computed at three scales: $\sigma_{B_1} = 1$, $\sigma_{B_2} = 2$, $\sigma_{B_3} = 8$ (Table 1). The optimal window width was found to be 15 and 100 pixels, for the mESC and hESC PCM datasets respectively. This difference was most likely due to the nature of the textures to discriminate between in both cases.

These optimal results were compared to those obtained by using different feature types and encoding methods (only the best results obtained for each condition after parameter values optimisation are reported). For both datasets, using single-scale BIFs as features resulted in a slight drop in performance. Likewise, using single BIFs values instead of local BIFs histograms resulted in marked decreases in mean LOOCV F-score of 15% and 51% for the mESC and hESC PCM datasets, respectively. Finally, alternative feature types were considered: raw pixel intensity values and pixel intensity values after application of a contrast (i.e. standard deviation) filter. Both feature types were encoded using 10 bin local histograms. The same scales permutations and window widths were investigated as in the case of BIFs.

Using intensity pixel values as features resulted in a pronounced decrease in segmentation performance for both datasets. In contrast, standard deviation features fared well, especially for the hESC datasets where it approached the performance of single-scale BIFs.

Our results were then compared with those obtained using previously described PCM image segmentation algorithms for the same mESC PCM images dataset (Table 1). Our trainable segmentation approach outperformed two of the three specialised algorithms it was compared to and produced results approaching those obtained using the third (best performing) one. It is important to note that those algorithms were specifically devised based on *a priori* knowledge of PCM images properties and specificities whereas trainable segmentation employed a generic framework solely based on BIFs and the user-set hard constraints. Interestingly, the classifier learned how to properly label halo artefacts around foreground objects as background pixels without being explicitly designed to do so (Fig.2).

Table 1: Optimal segmentation results compared with those obtained by alteration of various components of the algorithm: single BIFs scales in place of multiple scales, a single pixel value instead of local histograms, and intensity or contrast features instead of BIFs. Our results for the mESC dataset were also compared to specialised PCM segmentation algorithms. All results shown as mean F-score $\pm$ standard deviation after leave-one-out cross-validation (LOOCV).

| | mESC dataset | hESC dataset |
|---|---|---|
| Optimal | $0.92 \pm 0.05$ | $0.90 \pm 0.07$ |
| Single-scale scheme | $0.91 \pm 0.05$ | $0.88 \pm 0.10$ |
| No histogram | $0.78 \pm 0.15$ | $0.44 \pm 0.22$ |
| Intensity features | $0.83 \pm 0.12$ | $0.71 \pm 0.24$ |
| Contrast features | $0.85 \pm 0.15$ | $0.87 \pm 0.10$ |
| Jaccard et al [6] | $0.95 \pm 0.04$ | - |
| Juneau et al [7] | $0.85 \pm 0.10$ | - |
| Topman et al [14] | $0.84 \pm 0.11$ | - |

# 4 Summary and conclusion

In this work, we described a trainable segmentation algorithm for PCM images based on multi-scale local BIFs histograms. It performed well in foreground versus background segmentation tasks, approaching performance of state-of-the-art specialised algorithms. Indeed, the random forest classifier implicitly learned how to correct halo artefacts, which is usually done as an extra post-processing step in said algorithms [2, 6]. It also produced good results for a more complex segmentation task consisting in differentiating between two types of foreground objects with similar visual attributes. The fact that two significantly different problems could be suitably addressed using the same algorithm demonstrated the versatility of trainable segmentation approaches in general, and that of the proposed method in particular.

Processing a standard microscopy image ($1280 \times 960$ pixels) took less than 4 seconds using a single thread on an 3.7 Ghz E5-1620 CPU, including the computations of BIFs at three scales and the construction of histograms for each pixel of the image. Using BIFs as features had the advantage of requiring only 7 bin histograms per scale, which allowed their rapid computation for each pixel. It also significantly reduced the computational complexity

of the offline phase (i.e. classifier training) as memory requirements and training time both increase with the number of features. In contrast, when using local 256-bin intensity feature histograms, computation time soared to more than 45 seconds for the same image and conditions. Specialised algorithms took on average about a second to process the same images [6].

These low processing times using BIFs make this method suitable for batch segmentation of large number of PCM images or that of time-lapse movie frames. To generate the results presented in this paper, the classifier was trained based on 50000 pixels sampled across the entire dataset (minus the left out image), or less than 1.6% and 1% of the total number of pixels for the mESC and hESC datasets, respectively. Combined with the low processing times, the ability to handle sparse annotations could enable the use of the proposed approach for interactive segmentation of PCM images.

# 5 Acknowledgment

The authors gratefully acknowledge the British Heart Foundation for funding Nicolas Jaccard's PhD studentship (Grant no. SP/08/004)

# References

[1] M.E. Ambühl, C. Brepsant, J.-J. Meister, A.B. Verkhovsky, and I.F. Sbalzarini. High-resolution cell outline segmentation and tracking from phase-contrast microscopy images. *Journal of microscopy*, 245(2):161–70, 2012.

[2] C.J. Bradhurst, W Boles, and Y Xiao. Segmentation of bone marrow stromal cells in phase contrast microscopy images. In *23rd IVCNZ*, pages 1–6. IEEE, 2008.

[3] O. Debeir, I. Adanja, N. Warzee, P. Van Ham, and C. Decaestecker. Phase contrast image segmentation by weak watershed transform assembly. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 724–727. IEEE, 2008.

[4] L.D. Griffin, M. Lillholm, M. Crosier, and J. van Sande. Basic image features (bifs) arising from approximate symmetry type. In *Scale Space and Variational Methods in Computer Vision*, pages 343–355. Springer, 2009.

[5] L.D. Griffin, P. Elangovan, A. Mundell, and D.C. Hezel. Improved segmentation of meteorite micro-ct images using local histograms. *Computers & Geosciences*, 39:129–134, 2012.

[6] N. Jaccard, L.D. Griffin, A. Keser, R.J. Macown, A. Super, F.S. Veraitch, and N. Szita. Automated method for the rapid and precise estimation of adherent cell culture characteristics from phase contrast microscopy images. *Biotechnology and bioengineering*, 111 (3):504–17, 2014.

[7] P.-M. Juneau, A. Garnier, and C. Duchesne. Selection and tuning of a fast and simple phase-contrast microscopy image segmentation algorithm for measuring myoblast growth kinetics in an automated manner. *Microscopy and microanalysis*, 19(4):855–66, 2013.

[8] T. Kazmar, M. Smid, M. Fuchs, B. Luber, and J. Mattes. Learning cellular texture features in microscopic cancer cell images for automated cell-detection. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010:49–52, 2010.

[9] N OTSU. Threshold Selection Method From Gray-Level Histograms. *IEE Transactions on Systems Man and Cybernetics*, 9(1): 62–66, 1979.

[10] M. Reichen, R.J. Macown, N. Jaccard, A. Super, L. Ruban, L.D. Griffin, F.S. Veraitch, and N. Szita. Microfabricated Modular Scale-Down Device for Regenerative Medicine Process Development. *PLoS ONE*, 7(12):e52246, 2012.

[11] F. Schroff, A. Criminisi, and A. Zisserman. Object Class Segmentation using Random Forests. In *Procedings of the British Machine Vision Conference 2008*, pages 54.1–54.10, 2008.

[12] I. Seroussi, D. Veikherman, N. Ofer, S. Yehudai-Resheff, and K. Keren. Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes. *Journal of microscopy*, 247(2):137–46, 2012.

[13] C. Sommer, C. Straehle, U. Kothe, and F.A. Hamprecht. Ilastik: Interactive learning and segmentation toolkit. In *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 230–233. IEEE, 2011.

[14] G. Topman, O. Sharabani-Yosef, and A. Gefen. A Method for Quick, Low-Cost Automated Confluency Measurements. *Microscopy and Microanalysis*, 17(06):915–922, 2011.

[15] Z. Yin, R. Bise, M. Chen, and T. Kanade. Cell segmentation in microscopy imagery using a bag of local Bayesian classifiers. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 125–128. IEEE, 2010.

[16] Z. Yin, T. Kanade, and M. Chen. Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation. *Medical image analysis*, 16(5):1047–62, 2012.

# Detecting and Segmenting Nanodiscs in Immuno-Electron Micrographs

Tianjun Huang[1]
t.huang@dundee.ac.uk
Christian Hacker[2]
ch84@st-andrews.ac.uk
John Lucocq[2]
jml7@st-andrews.ac.uk
Stephen McKenna[1]
stephen@computing.dundee.ac.uk

[1] School of Computing
University of Dundee
Dundee, UK

[2] School of Medicine
University of St Andrews
St Andrews, UK

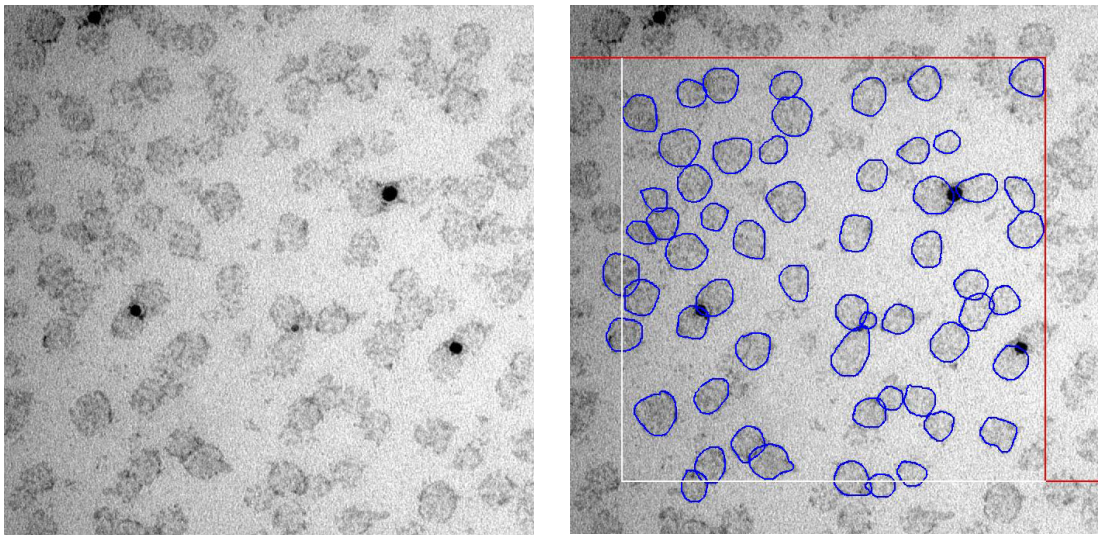
**Abstract**

Nanodiscs are soluble nanoscale phospholipid bilayers with applications in drug delivery and the study of membrane proteins, for example. They can be imaged using electron microscopy, along with immunogold markers indicating locations of proteins of interest. We describe and evaluate methods for automatically detecting and segmenting nanodiscs in electron micrographs. The detection method modifies aspects of the Fast Radial Symmetry Transform to detect nanodiscs that exhibit approximate radial symmetry against noisy but predominantly lighter background. Detected nanodiscs are then segmented using radial active contours. Experiments on micrographs both with and without immunogold markers indicate promising detection and segmentation performance, and that information on nanodisc quantities, locations, size distributions, and co-location with proteins of interest could be extracted automatically.

## 1 Introduction

Nanodiscs are soluble nanoscale phospholipid bilayers composed of a genetically engineered membrane scaffold protein and phospholipid [2]. They are useful for studying the function and structure of membrane proteins and have applications such as vehicles for transport of hydrophobic drugs [7]. Electron microscopy (EM) can be used to image nanodiscs, as well as to locate proteins via immunogold markers that appear as dark spots in the electron micrographs. This paper investigates automating the detection and segmentation of nanodiscs, in the presence of immunogold markers, to enable information such as nanodisc quantities, locations, size distributions, and co-location with immunogold-marked proteins to be measured efficiently during studies of protein structure and function or of applications such as drug delivery. We are not aware of any previously published literature describing how to automatically detect and segment nanodiscs using EM.

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.



(a) (b)

Figure 1: (a) Nanodisc image with gold markers. (b) Manual annotation using forbidden lines.

## 2 Detecting Nanodiscs

Figure 1 shows an example EM image containing nanodiscs as well as five immunogold markers. When imaged in this way, nanodiscs tend to be characterised by approximate radial symmetry, noisy interiors which are on average darker than the background, and darkening around their boundaries. We detect them using a method inspired by the Fast Radial Symmetry Transform (FRST) that was proposed by Loy and Zelinsky to detect interest points exhibiting local radial symmetry [5]. FRST has been used previously as a component in cell nuclei detection [8]. Ni *et al.* [6] extended FRST to cope with affine transformations and applied it to nuclei detection in histopathology. Here we modify FRST to detect nanodiscs.

The detector searches for nanodiscs at radii $n \in N$ where $N$ is a set of radii sampling a known range of nanodisc radii (in pixels). Image gradients are computed (using Farid and Simoncellis' algorithm [3]) and pixels at which the gradient magnitude exceeds a threshold $\tau_{mag}$ vote to produce an *orientation projection image* $O_n$ and a *magnitude projection image* $M_n$. Specifically, a pixel at location $\mathbf{p}$ votes for the pixel location nearest to $\mathbf{p} - n\frac{g(\mathbf{p})}{||g(\mathbf{p})||}$, where $g(\mathbf{p})$ is the gradient at $\mathbf{p}$. After voting is complete, the value of $O_n(\mathbf{p})$ is the number of pixels that voted for $\mathbf{p}$ and $M_n(\mathbf{p})$ is the sum of the gradient magnitudes of all the pixels that voted for $\mathbf{p}$. Similarly to [5] these are combined to give $G_n(\mathbf{p}) = M_n(\mathbf{p})(\min(k, O_n(\mathbf{p})))^2$ where $k$ was set to approximately 20. However, rather than average the response at different radii as in [5], we find the maximal response $S(\mathbf{p}) = \max_n G_n(\mathbf{p})$ and treat this as evidence for a nanodisc of radius $n^* = \arg\max_n G_n(\mathbf{p})$ centred at $\mathbf{p}$. A nanodisc will tend to have high response, $S$, near its centre but this response will be spread and can even be multimodal. We apply a local averaging filter to $S$ and then set response values not exceeding a detection threshold $\tau_{detect}$ to zero. Non-maximum suppression is then used to locate candidate nanodisc centres (with the radius of the suppression window set to the known minimum nanodisc radius).

Before running the nanodisc detector, any immunogold markers in the image are detected using the method described by Wang *et al.* [9]. Immunogold marker regions are then dilated

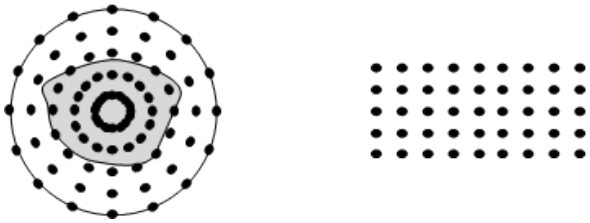


Figure 2: Left: Sampled annular search space. Right: Corresponding trellis.

and any pixels within these dilated regions are excluded from voting for nanodisc centres. This is important to avoid the presence of immunogold markers adversely affecting nanodisc detection. Impurities in the electron micrograph can result in very dark regions; these are ignored by simply removing any candidate nanodisc centre with an image intenisty below a conservative threshold $\varepsilon$. If the distance between the centres of two candidates is less than any one of the candidates' radii, the candidate which has higher symmetry response is kept. Finally, any remaining candidates are labelled as nanodiscs. In the experiments, we set $\tau_{mag}$ to 2% of the maximum gradient magnitude, $\tau_{detect} \approx 0.1$ and $\varepsilon = 20$.

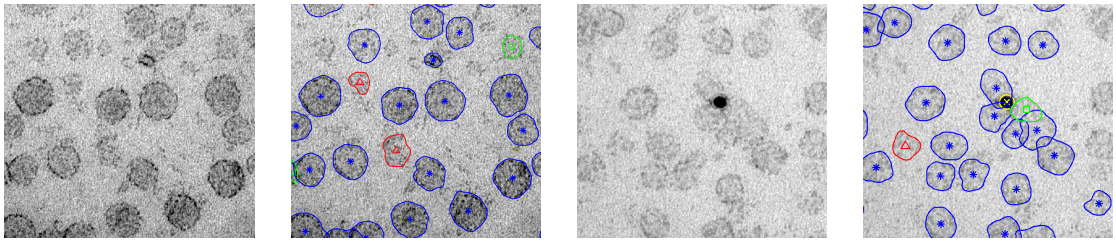
## 3 Segmenting Detected Nanodiscs

Each nanodisc detection is used to initialise a radial active contour segmentation similar to that used by Bamford and Lovell to segment cell nuclei [1]. The radius and the centre location estimated by the nanodisc detector are used to determine an annular search region within which the nanodisc's boundary contour is assumed to lie (see Figure 2). The annulus is sampled along $M$ radial lines at regular angular spacing, at $N$ equally spaced points on each radial line. The resulting search space can be thought of as an $N \times M$ trellis (right of Figure 2) and the search for the nanodisc contour posed as the search for an optimal left-right path across the trellis (where the left and right sides of the trellis are treated as adjacent). The Viterbi algorithm finds an optimal left-right path. Path optimality is with respect to a cost function that combines internal contour energy, $E_{int}$, and external energy, $E_{ext}$, as in Equation (1) where $\mathbf{v}_i$ represents a point in the $i^{th}$ column of the trellis. The internal energy is given in Equation (2); it encourages smooth contours and is minimised by circles.

$$C_i(\mathbf{v}_{i+1}, \mathbf{v}_i) = \min[C_{i-1}(\mathbf{v}_i, \mathbf{v}_{i-1}) + \lambda E_{int}(\mathbf{v}_{i-1}, \mathbf{v}_i, \mathbf{v}_{i+1}) + (1-\lambda)E_{ext}(\mathbf{v}_i)] \tag{1}$$

$$E_{int} = \frac{\|\mathbf{v}_{i-1} - 2\mathbf{v}_i + \mathbf{v}_{i+1}\|^2}{\|\mathbf{v}_{i-1} - \mathbf{v}_{i+1}\|^2} \tag{2}$$

The parameter $\lambda \in [0,1]$ balances internal and external energies. We set $\lambda = 0.7$. At each stage, the minimum cost at the current point and also its corresponding path from the point of the previous stage will be recorded. The final path is obtained by back-tracking from the point in the last column of the trellis. In order to obtain a closed contour we adopt a technique proposed by Gunn [4]; two points at the middle of the trellis on the open contour obtained by an initial Viterbi search are fixed as endpoints for a second Viterbi search. Furthermore, after each contour search, a new nanodisc centre can be estimated from the contour and used to determine a new annular search region. We iterated the search twice.



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

Figure 3: (a) Part of a nanodisc image without gold markers and (b) its result. (c) Part of a nanodisc image with gold markers and (d) its result. Blue: correctly detected nanodiscs. Red: falsely detected nanodisc. Green: missed nanodiscs. Yellow: gold marker detections.

Using the inverse of the gradient magnitude at $\mathbf{v}_i$ as the external energy $E_{ext}(\mathbf{v}_i)$ will encourage the contour to lie on image edges but will cause problems when a part of another nanodisc is included in the search space (as in Figure 5). Similarly to [1] $E_{ext}(\mathbf{v}_i)$ can be modified to instead use the component of the image gradient in the radial direction (*i.e.*, projected onto a vector from the hypothesised nanodisc centre to the image point being considered). This encourages the contour to lie on image edges that are approximately in the radial direction and which are consistent with an object that is darker than the background. This method is referred to as the *directional gradient* method.

Nanodiscs are sometimes quite far from circular and their contour directions can then sometimes by poorly approximated by the radial direction. This can result in the *directional gradient* method giving poor results. Therefore, we tried a third external energy which combines the gradient magnitude and the directional gradient:

$$E_{ext} = w|\nabla_{directional}| + (1-w)|\nabla| \tag{3}$$

where $|\nabla|$ is the gradient magnitude, $\nabla_{directional}$ is the directional gradient, and $w \in [0,1]$ is a weight parameter which we set as $w = 0.7$.

# 4 Experiments

Nanodiscs and immunogold markers were manually annotated by C. Hacker in a data set consisting of five nanodisc images with immunogold markers (dataset A) and five without immunogold markers (dataset B). Images were $1024 \times 1024$ pixels and forbidden lines were placed at a distance of 10% of image width from the image border (see Figure 1). Dataset A had 304 nanodiscs and 15 immunogold markers annotated. Dataset B had 199 nanodiscs. Manual annotation took approximately 12 minutes per image in dataset A and 7 minutes per image in dataset B. A Matlab implementation of the proposed method took approximately 90*s* per image.

Figure 3 shows an example cropped result for each dataset. The curves in Figure 4, obtained by varying the symmetry response threshold, $\tau_{detect}$, show false positive and false negative nanodisc detection rates per 100 annotated nanodiscs. The segmentation of each correctly detected nanodisc was compared to its manual segmentation using the Jaccard Index, $J = \frac{A_a \cap A_s}{A_a \cup A_s}$, where $A_a$ and $A_s$ are the sets of pixels in the manual and automatic segmentations respectively. Table 1 compares the results of the proposed segmentation method with

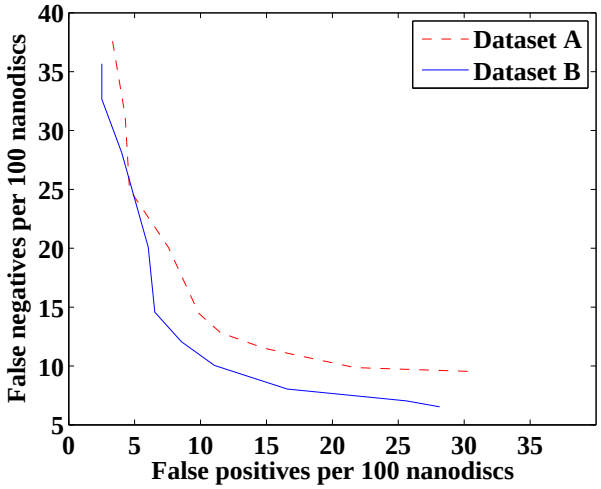



Figure 4: False positives and false negatives per 100 nanodiscs.

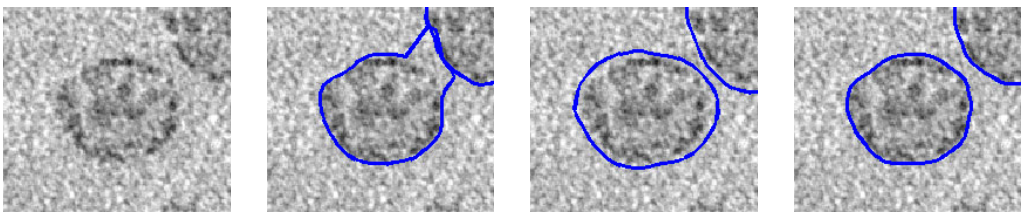
| | Dataset A | Dataset B |
|---|---|---|
| Gradient magnitude | 0.74 (0.12) | 0.75 (0.10) |
| Directional gradient | 0.69 (0.12) | 0.69 (0.10) |
| Gradient magnitude + Directional gradient | 0.76 (0.12) | 0.77 (0.09) |

Table 1: Mean Jaccard indices when comparing automatic and manual nanodisc segmentations (standard deviations in parentheses).

the method using only magnitude or directional gradient in the external energy. The proposed method compares favourably and Figure 5 shows an example where this is apparent. Adjacent nanodiscs can cause difficulties in distinguishing the contour in the overlapping regions; however, reasonable results can still be obtained as shown in Figure 6. All immunogold markers were detected successfully.

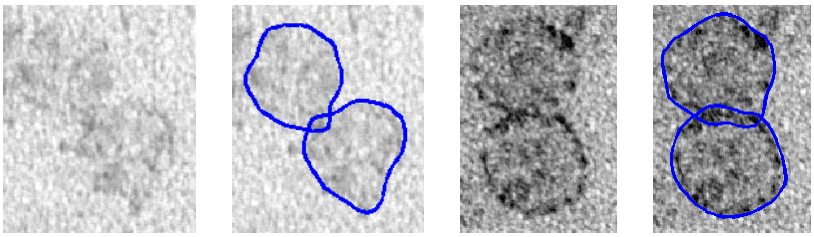
# 5 Conclusion

The experiments show that nanodiscs can by detected and segmented with reasonably high accuracy, including in the presence of immunogold markers. Detection failures occurred in



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

Figure 5: (a) Example of cropped input, (b-d) results using gradient magnitude method, directional gradient method, and proposed combined method respectively.



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

Figure 6: Segmentation of adjacent nanodiscs: (b) and (d) are the segmentations of (a) and (c) respectively.

cases where it was difficult to distinguish nanodiscs from background noise even for experts. We are currently investigating inter-observer variability. Two observers were asked to mark independently all the nanodiscs in a set of 10 images (five with and five without immunogold markers). Preliminary data shows 528 nanodiscs were marked by at least one of the two observers. Of these, 454 were marked by both observers. This gives a Jaccard Index of $454/528 \approx 0.86$. Taken in this context, the detection results reported in Section 4 are promising.

# References

[1] P. Bamford and B. Lovell. Unsupervised cell nucleus segmentation with active contours. *Signal Processing*, 71(2):203 – 213, 1998.

[2] T. H. Bayburt and S. G. Sligar. Membrane protein assembly into nanodiscs. *FEBS Letters*, 584:1721–1727, 2010.

[3] H. Farid and E. P. Simoncelli. Differentiation of discrete multidimensional signals. *Image Processing, IEEE Transactions on*, 13:496–508, 2004.

[4] S. R. Gunn. *Dual active contour models for image feature extraction*. PhD thesis, University of Southampton, 1996.

[5] G. Loy and A. Zelinsky. Fast radial symmetry for detecting points of interest. *PAMI, IEEE Transactions on*, 25:959–973, 2003.

[6] J. Ni, M. K. Singh, and C. Bahlmann. Fast radial symmetry detection under affine transformations. In *CVPR*, pages 932–939, 2012.

[7] R. O. Ryan. Nanodisks: hydrophobic drug delivery vehicles. *Expert Opinion on Drug Delivery*, 5(3):343–351, 2008.

[8] N. Timilsina, C. Moffatt, and K. Okada. Development of a stained cell nuclei counting system. In *Medical Imaging: Image Processing*, volume Proc. SPIE Vol. 7962, pages 79620K–1, 2011.

[9] R. Wang, H. Pokhariya, S. J. McKenna, and J. Lucocq. Recognition of immunogold markers in electron micrographs. *Journal of Structural Biology*, 176:151–158, 2011.

# Finding Golgi Stacks in Electron Micrographs

Neil Fordyce[1]
neilfordyce@hotmail.com
Stephen McKenna[1]
stephen@computing.dundee.ac.uk
Christian Hacker[2]
ch84@st-andrews.ac.uk
John Lucocq[2]
jml7@st-andrews.ac.uk

[1] School of Computing
University of Dundee
Dundee, UK

[2] School of Medicine
University of St Andrews
St Andrews, UK

**Abstract**

Transmission electron microscopy (EM) can acquire images in which a range of subcellular organelles are clearly resolved simultaneously. There exist mature stereology techniques for extracting quantitative specimen information from section-based EM images and such techniques have been adopted successfully for use in immuno-EM. A bottleneck preventing the application of these nanomorphomics methods to high throughput applications is the recognition of organelle structures. This papers addresses this issue for one important organelle, the Golgi apparatus. A support vector machine is trained as a local Golgi detector based on rotationally invariant features. The SVM output is used to drive a graph-cuts segmentation. The ability of the method to detect and segment Golgi stacks is evaluated on a set of 36 micrographs.

## 1 Introduction

Transmission electron microscopy (EM) can acquire images in which a range of organelles and their substructures are clearly resolved simultaneously. Furthermore, there exist mature stereology techniques for extracting quantitative specimen information from section-based EM images, providing unbiased and efficient estimates of 3D volumes, surfaces, and numbers as well as the distributions of organelles (e.g. [6]). Such techniques have been adopted successfully for use in immuno-EM, a method that maps molecular components and their concentrations over a range of subcellular compartments using particulate nanogold markers [10]; detection of the nanogold markers can be automated [16]. A bottleneck preventing the application of these *nanomorphomics* methods to high throughput data mining is the recognition of organelle structures in the images, a task which is carried out by trained experts. This paper addresses the automatic recognition and segmentation of one important organelle, the Golgi apparatus. The Golgi apparatus includes a stack of flat membrane-bound structures called cisternae. This stack usually presents visually in section as repeating curvilinear structures. Vesicles can often be seen in proximity to the stack. Figure 1(a) shows

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

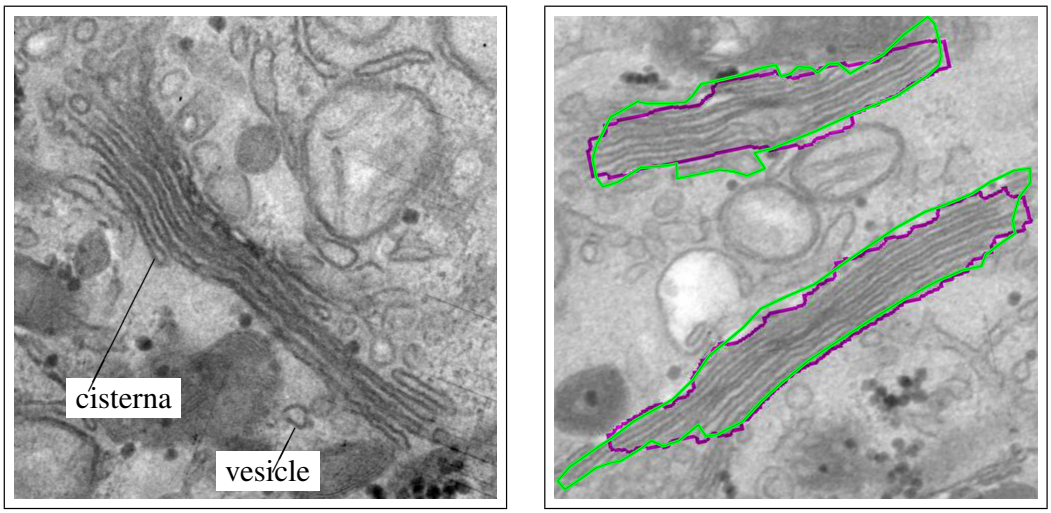



Figure 1: EM micrograph images cropped to 2000x2000 pixels. (a) Golgi apparatus with its stack of cisternae. (b) Manual (green) and automatic (magenta) segmentations.

an electron micrograph in which Golgi apparatus is clearly visible. In Figure 1(b), manual annotations of the extent of Golgi stacks are overlaid, along with a segmentation obtained automatically using the method described in this paper.

## 2 Related Work

We are not aware of any previous literature on automatic recognition of Golgi apparatus in EM images. There is however published work regarding other organelles and structures. For example, progress has been made on segmentation of mitochondria [2, 12, 13]. Perhaps most promisingly, Lucchi et al. [8] segmented mitochondria from surrounding background using support vector machine classifiers to assign probabilities to superpixels based on local texture and shape descriptors that were then used in a graph cuts optimization. A similar approach was applied to FIBSEM image stacks [9].

Regarding other structures, Nam et al. [11] segmented secretory granules using level set active contours and membrane sampling. Kreshuk et al. [5] detected synapses in intact nervous tissue in a FIBSEM image by training a random forest classifier to classify voxels based on local features. Others have presented methods for automatic enhancement and localisation of thin elongated structures in EM images [4, 14, 15].

## 3 Method

Firstly, a detection stage uses an SVM detector trained on subwindows of EM images. The detector uses features based on Histogram of Oriented Gradients (HOG). These features extend standard HOG to provide rotation-invariant Fourier HOG [7] descriptors. The key idea is treating gradient histograms as continuous functions, analysing polar coordinate gradient images in Fourier space. The features are the same regardless of the orientation of the Golgi stack which is arbitrary. Due to noise from the image acquisition technique, robustness to

noise is also desirable. Fourier HOG features offer both of these characteristics. The first three degrees of Fouier basis are used for the Fourier representation. The feature vector for each window comprises of 78 real-valued features. A binary SVM is trained on features extracted from hand annotated images, to discriminate between Golgi and non-Golgi subwindows. Each sampled feature vector describes a subwindow of the image. Using a sliding window approach, the SVM is used to obtain an SVM score image which forms the output of the detection stage.

Secondly, graph cuts is used to produce segmentations by combining the detector output with spatial information from the original EM image. We use a graph cuts library [1] which implements approximate energy minimisation techniques. As in [3], our energy function, $E$, of a labelling, $L$, can be expressed as

$$E(L) = \sum D_p(L_p) + \lambda \sum V_{p,q}(L_p, L_q) \tag{1}$$

where $\lambda$ is a parameter that determines the relative importance of a spatially coherent segmentation to one which is consistent with the results of the detection stage. $D$ is a data cost w.r.t. the SVM probability image. It penalises assignment of pixels to labels with SVM scores inconsistent with other SVM scores in that labelling. Specifically, the data cost of assigning a label $L_p$ to a pixel with SVM score $s_p$, is given by Equation (2) where $\mu_L$ and $\sigma_L^2$ are the mean and variance, respectively, of SVM scores of pixels assigned to label $L$.

$$D_p(L_p) = \frac{(s_p - \mu_L)^2}{\sigma_L^2} \tag{2}$$

$V$ penalises crossing gradient boundaries in the EM image and encourages a smooth final segmentation which adheres to image contours. For adjacent pairs of pixels $I_p$ and $I_q$, labelled $L_p$ and $L_q$ respectively, we have:

$$V_{p,q}(L_p, L_q) = \begin{cases} exp(-|I_p - I_q|) & \text{if } L_p \neq L_q \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

As a final step, any segmented components with an area below a threshold are removed.

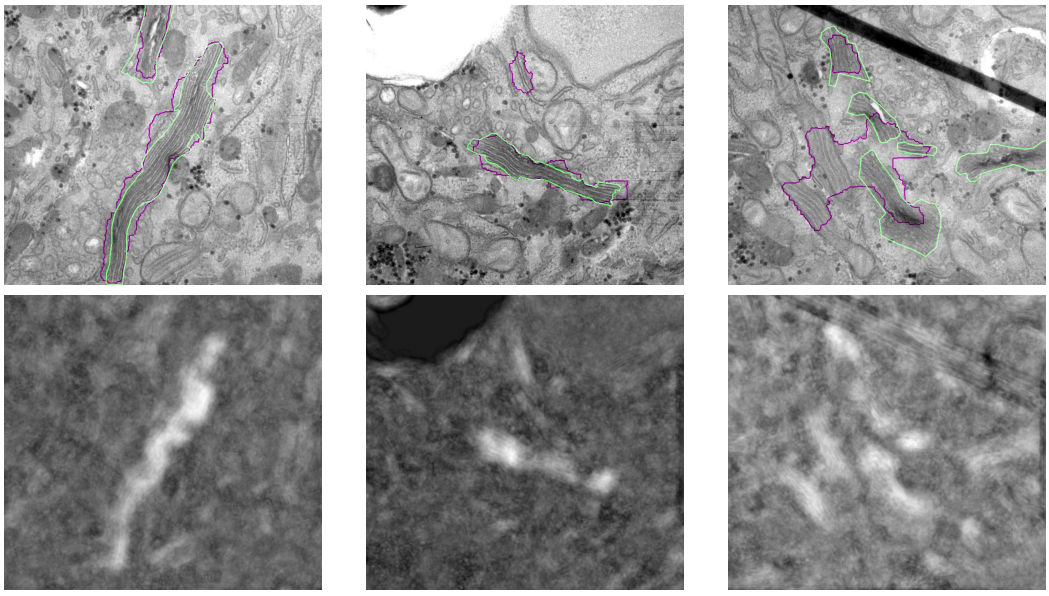
# 4 Experiments

Our data set consisted of 36 electron micrographs of rabbit kidney cells. Each image had a resolution of $4725 \times 4167$ pixels. Images were subsampled to $945 \times 834$ pixels for feature extraction and the detector window size was then $31 \times 31$ pixels.

The Golgi stacks in the dataset were of various shapes and sizes. Some were very clearly defined and in focus, while other stacks appeared blurred and were identifiable in part due to proximal vesicles. All Golgi stacks were segmented manually to provide ground truth. Golgi which were badly blurred were given an additional annotation of *blurred*. Two sets of results are presented: one in which all Golgi were required to be detected, the other in which blurred Golgi were discounted.

Five-fold cross-validation was used to train and test detectors. The Fourier HOG features create densely populated descriptors for each image. 7200 negative and 9600 positive training examples were extracted at random from the training images to train an SVM.

Figure 2 shows example Golgi images with automatic segmentations overlaid, as well as the corresponding SVM outputs. Figure 2(b) shows two segmentation attempts: one of



(a) (b) (c)

Figure 2: Top row: manual (green) and automatic (magenta) segmentations of Golgi stacks. Bottom row: SVM detector score images used as input to the graph cuts segmentation.

these is a reasonable segmentation, while the other is a false segmentation caused by the membranes of several other organelles being aligned to appear similar to Golgi. Figure 2(c) shows two segmented regions: one of these is reasonable, the other contains parts of three separate Golgi stacks and nearby mitochondria. The mitochondria in the segmentation have an appearance resembling Golgi cisternae. The proximity of multiple Golgi and mitochondria have brought about a single large segmentation. The blurred appearance of another Golgi in 2(c) has caused it to be missed.

Figure 3 shows an ROC curve obtained by varying the SVM score threshold. The F1 score (F-measure) was 0.3157 (0.3235 when blurred Golgi were excluded); this low score arose from a high false positive rate; however, the false positives are often spatially isolated so are subsequently removed by graph cuts. The optimal pixel misclassification rate obtained was 0.1589 and this reduced to 0.0983 when blurred Golgi regions were excluded.

A Jaccard index was used to quantify the quality of segmentation as follows. A binary labelling $L$ and binary ground truth $G$ can be combined to produce an evaluation image $I = G + 2L$. A Jaccard index is obtained for each connected region $R$ of non-zero pixels in $I$. So $p \in R \Rightarrow I(p) > 0$. Let $A = \{p | p \in R, I(p) = 3\}$ and let $B = \{p | p \in R, I(p) \geq 1\}$. Then the Jaccard index is defined for each $R$ as, $J = \frac{|A \cap B|}{|A \cup B|}$. If multiple distinct Golgi regions are labelled as a single region, then the regions are evaluated as a single segmentation. If a single Golgi region is labelled with multiple segmentations, then the region is evaluated as a single segmentation. A schematic of this is shown in Figure 4; although one labelling contains multiple Golgi and one Golgi is labelled in parts, the entire region (red, green and blue) is counted as a single segmentation and assigned a single Jaccard index. If a Golgi region does not intersect with any region labelled as Golgi, this is counted as a missed segmentation. On the other hand, if a region labelled as Golgi does not intersect with any Golgi, this is counted as a false segmentation.

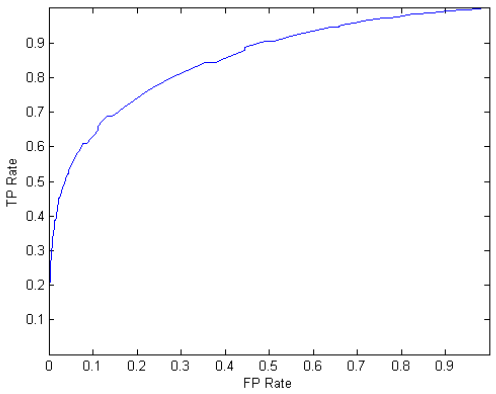



Figure 3: ROC curve (pixel classification)

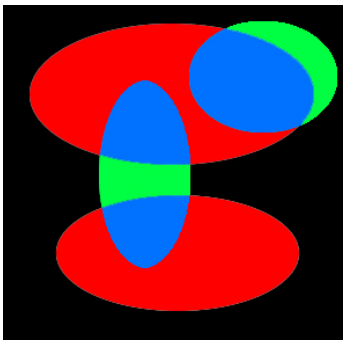


Figure 4: Segmentation schematic: Red=missed Golgi, green=incorrectly labelled Golgi, blue=correctly labelled Golgi, black=correctly labelled non-Golgi

| | Inc. blur | Exc. blur |
|---|---|---|
| True detections | 40 | 35 |
| False detections | 10 | 10 |
| Missed Golgi | 20 | 5 |

Table 1: Golgi detection results (including and excluding blurred Golgi regions)

Table 1 summarises segmentation results (with and without blurred regions). Although the number of missed Golgi was quite high, the majority of these came from blurred regions. The number of true detections dropped when blurred regions were excluded, indicating that some blurred Golgi had been detected. The mean Jaccard index across the true detections was 0.58 when blurred Golgi were included and 0.63 when they were excluded.

# 5 Conclusion

This paper contributes to the goal of automating organelle recognition in EM by demonstrating automatic detection and segmentation of Golgi stacks. High scoring segmentations occured when Golgi stacks were in focus and some distance from other organelles. False segmentations largely arose from mitochondria and aligned membranes which appeared similar to Golgi. The majority of missed segmentations were of Golgi which were badly blurred. Considering vesicles associated with the Golgi stack could improve performance in these areas.

# References

[1] S. Bagon. Matlab wrapper for graph cuts, December 2006. URL http://www.wisdom.weizmann.ac.il/~bagon.

[2] C. Bazán, M. Miller, and P. Blomgren. Structure enhancement diffusion and contour

extraction for electron tomography of mitochondria. *J. Structural Biology*, 166(2): 144–155, 2009.

[3] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. PAMI*, 26(9):1124–1137, 2004.

[4] M. Jiang, Q. Ji, and B. F. McEwen. Automated extraction of fine features of kinetochore microtubules and plus-ends from electron tomography volume. *IEEE Transactions on Image Processing*, 15(7):2035–2048, 2006.

[5] A. Kreshuk, C. N. Straehle, C. Sommer, U. Koethe, M. Cantoni, G. Knott, and F. A. Hamprecht. Automated detection and segmentation of synaptic contacts in nearly isotropic serial electron microscopy images. *PLoS One*, 6(10):e24899, 2011.

[6] S. L. B. Kristiansen and J. R. Nyengaard. Digital stereology in neuropathology. *APMIS*, 120(4):327–340, 2012.

[7] K. Liu, H. Skibbe, T. Schmidt, T. Blein, K. Palme, T. Brox, and O. Ronneberger. Rotation-invariant HOG descriptors using Fourier analysis in polar and spherical coordinates. *Int. J. Computer Vision*, 106(3):342–364, 2014.

[8] A. Lucchi, K. Smith, R. Achanta, V. Lepetit, and P. Fua. A fully automated approach to segmentation of irregularly shaped cellular structures in EM images. In *Proc. MICCAI*, volume LNCS 6362, 2010.

[9] A. Lucchi, K. Smith, R. Achanta, G. Knott, and P. Fua. Supervoxel-based segmentation of mitochondria in EM image stacks with learned shape features. *IEEE Transactions on Medical Imaging*, 30(11), 2011.

[10] J. Lucocq. Quantification of structures and gold labeling in transmission electron microscopy. *Methods in Cell Biology*, 88:59, 2008.

[11] D. Nam, J. Mantell, D. Bull, P. Verkade, and A. Achim. A novel framework for segmentation of secretory granules in electron micrographs. *Medical Image Analysis*, 18 (2):411 – 424, 2014.

[12] R. Narasimha, H. Ouyang, A. Gray, S. McLaughlin, and S. Subramaniam. Automatic joint classification and segmentation of whole cell 3D images. *Pattern Recognition*, 42 (6):1067 – 1079, 2009.

[13] H. Nguyen and Q. Ji. Shape-driven three-dimensional watersnake segmentation of biological membranes in electron tomography. *IEEE Trans. Medical Imaging*, 27(5): 616–628, 2008.

[14] D. Nurgaliev, T. Gatanov, and D. J. Needleman. Automated identification of microtubules in cellular electron tomography. In *Microtubules: in vivo*, volume 97 of *Methods in Cell Biology*, pages 475 – 495. Acad. Press, 2010.

[15] K. Sandberg and M. Brega. Segmentation of thin structures in electron micrographs using orientation fields. *J. Structural Biology*, 157(2):403 – 415, 2007.

[16] R. Wang, H. Pokhariya, S. J. McKenna, and J. Lucocq. Recognition of immunogold markers in electron micrographs. *J. Structural Biology*, 176:151–158, 2011.

# Cancer Imaging

# Analysing the Surface Morphology of Colorectal Polyps: Differential Geometry and Pit Pattern Prediction

Jingjing Zhang[1]
ZhangJ68@cardiff.ac.uk
Stephen J. McKenna[1]
stephen@computing.dundee.ac.uk
Jianguo Zhang[1]
jgzhang@computing.dundee.ac.uk
Maria Coats[2]
mariacoats@nhs.net
Frank A. Carey[3]
frank.carey@nhs.net

[1] CVIP, School of Computing,
University of Dundee,
Dundee DD1 4HN, UK
[2] School of Medicine,
Ninewells Hospital and Medical School,
Dundee DD1 9SY, UK
[3] Department of Pathology,
Ninewells Hospital and Medical School,
Dundee DD1 9SY, UK

**Abstract**

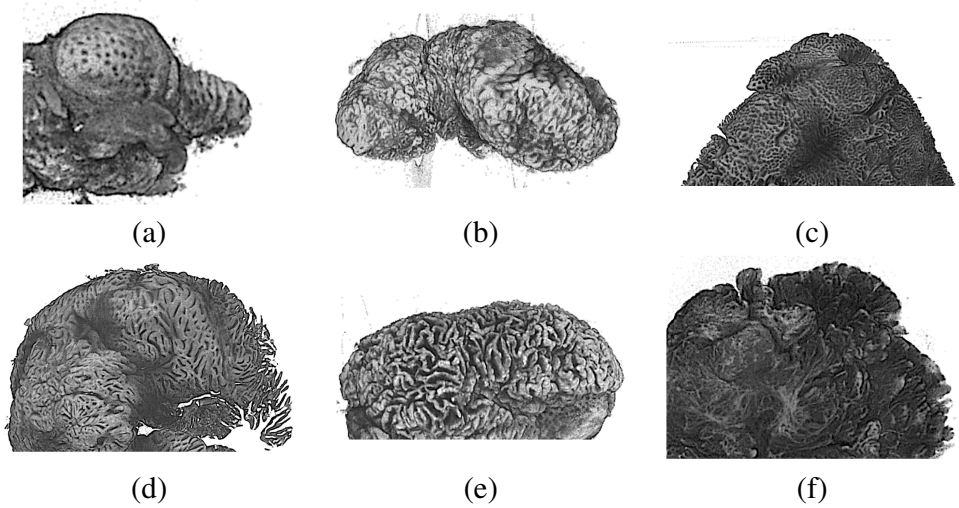
We present an initial study analysing the surface morphology of colorectal polyps from optical projection tomography. The differential geometry of polyp surfaces, segmented using a level sets method, is explored in terms of local, multi-scale shape index and curvedness descriptors. A surface region of interest can be represented using histograms of these descriptors. An experiment is described investigating the ability to predict pit pattern categories from these histograms using support vector machines.

## 1 Introduction

Colorectal cancer is the third most common cancer in men (746k cases, 10.0%) and the second in women (614k cases, 9.2%) worldwide [3]. Screening has reduced mortality by up to 21% and detected large numbers of adenomas and polypoid cancers [7]. However, analysis of histological H&E sections suffers from marked inter-observer variation.

There is a long history of study of polyp surface morphology relating to different stages of cancer development in the medical literature. Polyp surface morphology is often categorised in the pathology lab as being villous (having cerebral-like folds and finger-like protrusions), tubulo-villous or tubular (a smooth surface with pits/tubes in it). Whilst villous polyps are strongly associated with invasive cancer, all three types can and do develop into polypoid cancers[8]. This classification scheme is a coarse categorization of the complex and highly variable surface morphology and normally presents large inter-observer variation. Another categorization, used for in vivo endoscopic assessment based on visual appearance, is Kudo's pit patterns [5]. There are six classes of pit pattern: I, II, III-S, III-L, IV, and V.

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.



(a) (b) (c)

(d) (e) (f)

Figure 1: (a)-(f) OPT images with pit patterns I, II, III-S, III-L, IV and V respectively.

Here we investigate polyp morphology using optical projection tomography (OPT), an in vitro imaging technology capable of producing high-resolution 3D images of small biological specimens [11]. Li et al. [6] reported classification of 3D regions of dysplasia and invasive cancer in colorectal polyps from OPT. Here we analyse polyps' surface morphology. Figure 1 shows surface renderings of some OPT polyp images, one for each pit pattern.

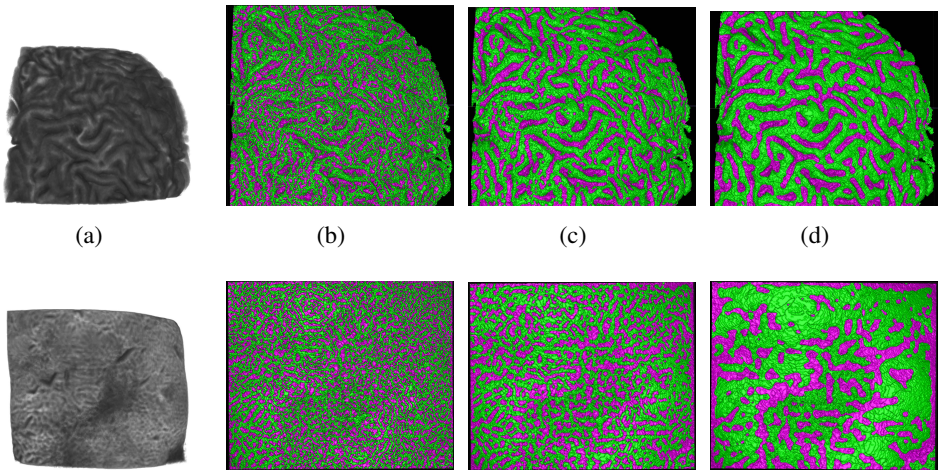
## 2 Surface Segmentation and Multi-scale Features

We employ a level sets fast marching method for polyp surface segmentation [10] using the implementation in [9]. Segmentation was seeded using 100 voxels randomly selected on a plane within the background. Once the surface is segmented, shape descriptors are constructed based on multi-scale differential geometry features computed at each voxel on the segmented surface. We do not fit a mesh grid which makes assumptions about topology. Instead, we work directly with image data and extract surface curvature estimates based on a 3D local neighbourhood in a similar fashion to [1]. This has the advantage of avoiding meshing errors. The principal curvatures $\kappa_{max}$ and $\kappa_{min}$ (maximum and minimum values of curvatures over all of directions) are obtained at each such voxel at several scales obtained by setting the Gaussian smoothing parameter, $\sigma$, used in the estimation of image gradients.

Figure 2 shows polyp surface regions rendered according to whether curvature is convex (green) or concave (magenta) at three scales. As $\sigma$ increases, the connected regions of concave and convex surface become larger and the coarse morphology becomes apparent. The characteristic scales exhibited by the different pit patterns are not all the same. For example, the surface morphology of the polyp region in Fig. 2(a) is perhaps best characterised at a larger scale ($\sigma = 4$) while the morphology of the region in Fig. 2(e) is perhaps best characterised at a smaller scale (e.g. $\sigma = 1$). This observation motivates a multi-scale analysis.

Rather than representing surface curvature directly in terms of principal curvatures $\kappa_{max}$ and $\kappa_{min}$, we compute a local *shape index*, $S$, and *curvedness*, $C$, as proposed by Koenderink [4] (Equations (1) and (2)).

$$S = -\frac{2}{\pi} \arctan \frac{\kappa_{max} + \kappa_{min}}{\kappa_{max} - \kappa_{min}} \tag{1}$$



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

(e) (f) (g) (h)

Figure 2: Two colorectal polyp surfaces classified as concave (magenta) and convex (green) at three scales. (a) A pit pattern IV region. (b-d) $\sigma = 1,2,4$ respectively. (e) A pit pattern III-S region. (f)-(h) $\sigma = 1,2,4$ respectively.

$$C = \frac{2}{\pi} \ln \left( \sqrt{\frac{\kappa_{max}^2 + \kappa_{min}^2}{2}} \right) \tag{2}$$

The shape index is a convenient measure of "which" shape, and the curvedness of "how much" shape. In the space spanned by $(C,S)$, all shapes are mapped onto a strip of infinite length in the $C$- direction. As the value of the shape index $S$ varies from $-1$ to 1, the corresponding surface shape transforms from concave pit, through trough, saddle, and ridge to convex peak. Figure 3 shows histograms of the shape index and curvedness obtained from surface regions of interest on eight polyps (two from each of pit patterns III-S, III-L, IV, and V). This gives some indication of the inter- and intra-class variability of these descriptors.

# 3 Pit Pattern Assignment

Experiments described in what follows used 28 OPT polyp images from 28 patients acquired using ultraviolet light and Cy3 dye. The size of each image was $1024 \times 1024 \times 1024$ voxels with aspect ratio of $1:1:1$. We included only polyps initially considered to be adenomatous polyps (pit patterns III-S, III-L, IV and V). Seven images were selected for each of the four pit patterns considered. The pit pattern type of each of these polyps was assigned by a trained pathologist at which time no secondary pit patterns were identified in these polyps. Their OPT images included some normal tissue and stalks; as such regions were not relevant for this study, they were excluded: a largest region of interest was selected through visual examination of each polyp volumetric image. The size of the extracted regions varied from $50 \times 300 \times 250$ to $350 \times 400 \times 500$.

As evidence of the reliability level of the assigned pit pattern labels, the same pathologist was asked at a later date to assign pit pattern labels to the 28 extracted regions. The

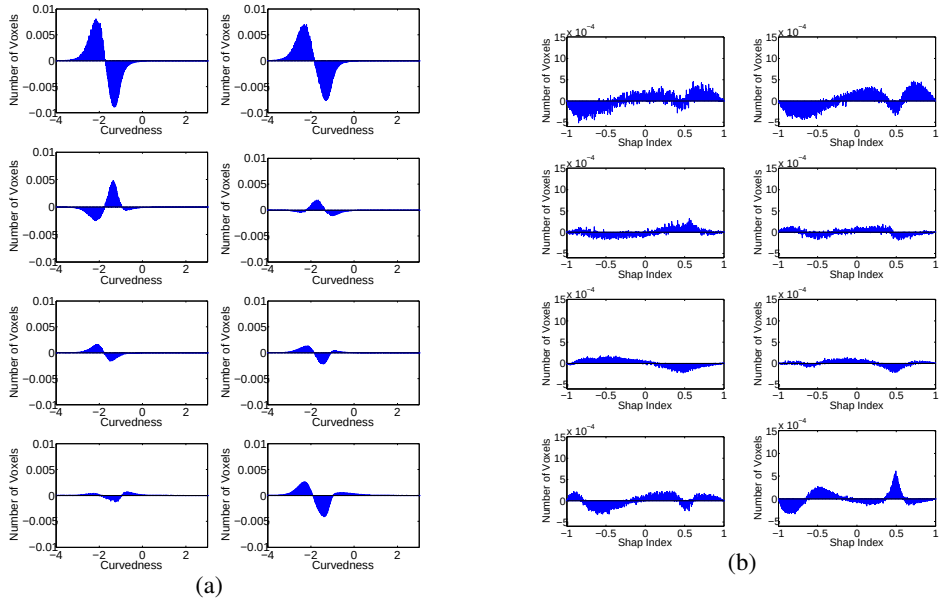



Figure 3: Histograms of (a) curvedness and (b) shape index from regions of interest on eight polyps. Each row represents one pit pattern (from top to bottom: III-S, III-L, IV and V). In each row, the first and third images are from the same polyp, and the second and the fourth are from another polyp. (Histograms have 1024 bins).

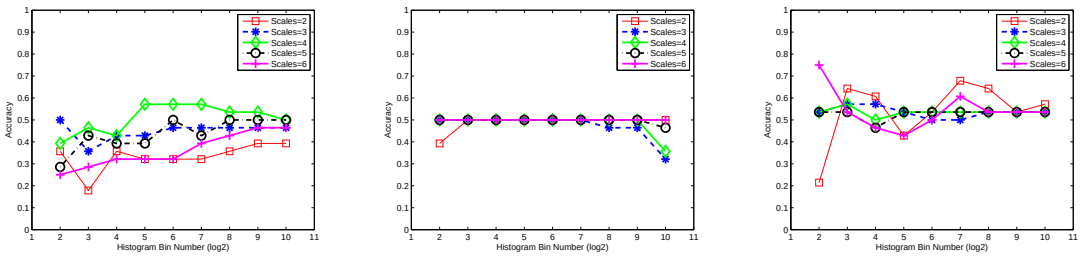
Table 1: Contingency table for pit pattern assignment by one observer. Initial session (whole polyp) vs. subsequent session (extracted region)

| | III-S | III-L | IV | V |
|---|---|---|---|---|
| III-S | **3** | 0 | 0 | 0 |
| III-L | 2 | **3** | 1 | 0 |
| IV | 2 | 2 | **6** | 4 |
| V | 0 | 0 | 0 | **3** |
| I | 0 | 2 | 0 | 0 |

regions were inspected in randomised order and the pathologist was blinded to any diagnostic information and to the previously assigned pit pattern labels. It was expected that the region extraction would have minimal effect on the pit pattern categorsiation, particularly as no secondary patterns had been noted. Nevertheless, the pit pattern assignment exhibited intra-observer variation; Table 1 is the resulting contingency table. The column represents the initial label assigned to the whole polyp while the row refers to the label subsequently assigned to the region. It can be seen that annotations of pit pattern IV are consistent over 6 samples, but for pit patterns III-S, III-L and V, only 3 in each category are consistent. In the region annotation, half of all the samples were categorized as pit pattern IV. In total, 15 polyps (54%) were annotated with the same label in both sessions.

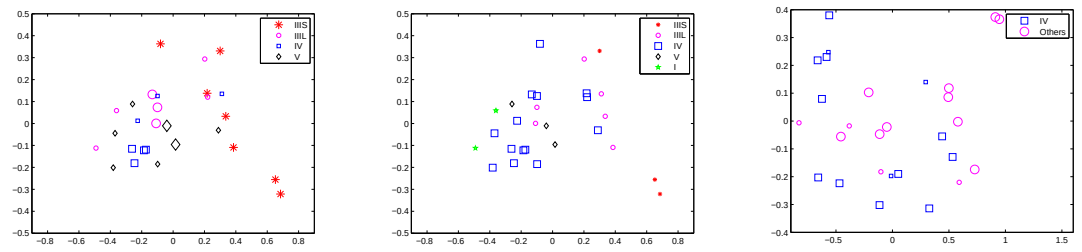
# 4 Pit Pattern Prediction Experiments

Support vector machines were trained to predict pit pattern labels. They were implemented using LIBSVM [2] and the 1-vs-rest method. As the number of polyps is relatively small,




(a) (b) (c)

Figure 4: SVM classification accuracy when predicting labels from (a) first session, (b) second session, and (c) in binary setting (IV vs. rest)




(a) (b) (c)

Figure 5: PCA visualisations. Correctly classified samples are larger. (a) First session labels, (b) second session labels, (c) binary labels

we report leave-one-out (LOO) cross validation results. To investigate the effect of scale we extracted features at scales of $\sigma = 2^k, k = 0,1,2,3,4,5$. To investigate the effect of the number of histogram bins, we used $2^k$ bins for $k$ ranging from 2 to 10. Thus the smallest number of bins was 4, and the largest was 1024.

Figure 4(a) shows accuracy at predicting pit pattern labels assigned in the first session. An accuracy of 57% was achieved with 32 bins and 4 scales ($\sigma = 1,2,4,8$). In this case, all 7 samples in pit pattern III-S were correctly classified while only 2 samples in pit pattern III-L, 4 samples in pit pattern IV and 2 samples in pit pattern V were correctly classified. This performance is in line with the variability of the pathologist (see Table 1).

We also trained SVMs to predict the pit pattern labels assigned by the pathologist in the second session; the pathologist had assigned the 28 samples to five pit patterns III-S, III-L, IV, V and I with 3, 6, 14, 3, 2 samples respectively (Table 1), resulting in an imbalanced dataset. SVM classification results under this setting are reported in Fig. 4(b). The highest classification accuracy obtained was 50% in which case all the samples of pit pattern IV were classified in agreement with the pathologist.

Finally, SVMs were trained on labels from the second session in a balanced, binary setting: Pit pattern IV vs other patterns (14 samples per class). Figure 4(c) shows the classification results; 75% accuracy was obtained (4 bins and $\sigma = 1,2,4,8,16,32$). Three samples in pit pattern IV and four samples in non-pit pattern IV were misclassified.

Figure 5 visualises the distribution of the samples in feature space after projection onto two dimensions using PCA. Correctly classified samples are depicted with markers of larger size than the incorrectly classified samples.

# 5 Conclusions and Future Work

We presented a preliminary investigation of the use of differential geometry features (shape index and curvedness) to characterise the surface morphology of colorectal polyps from OPT images. In particular we evaluated the ability to predict a pathologist's pit pattern assignments based on distributions of these features. It was noted during label assignment that some patterns observed may not fit well into Kudo's pit pattern categorisation. Results were broadly in line with intra-observer disagreement although further experiments using larger data sets are needed to confirm this and to understand the extent to which pit patterns capture clinically relevant surface morphological characteristics.

# References

[1] Brian Avants and James Gee. The shape operator for differential analysis of images. In *Information Processing in Medical Imaging*, pages 101–113. Springer, 2003.

[2] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[3] J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, and F. Bray. Globocan 2012 v1.0, cancer incidence and mortality worldwide: Iarc cancerbase no. 11 [internet]., 2013. URL http://globocan.iarc.fr. accessed on 25/03/2014.

[4] Jan J. Koenderink. *Solid Shape*, volume 2. Cambridge University Press, 1990.

[5] S. Kudo, S. Hirota, T. Nakajima, S. Hosobe, H. Kusaka, T. Kobayashi, M. Himori, and A. Yagyuu. Colorectal tumours and pit pattern. *J. Clin. Path.*, 47(10):880–885, Oct 1994.

[6] W. Li, J. Zhang, S. J. McKenna, M. Coats, and F. A. Carey. Classification of colorectal polyp regions in optical projection tomography. In *ISBI*, pages 736–739. IEEE, 2013.

[7] J. S. Mandel, T. R. Church, F. Ederer, and J. H. Bond. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. *Journal of the National Cancer Institute*, 91(5):434–437, Mar 3 1999.

[8] G. Nusko, U. Mansmann, U. Partzsch, A. Altendorf-Hofmann, H. Groitl, C. Wittekind, C. Ell, and EG Hahn. Invasive carcinoma in colorectal adenomas: multivariate analysis of patient and adenoma characteristics. *Endoscopy*, 29(07):626–631, 1997.

[9] Gabriel Peyre. Toolbox fast marching (matlab software), 2009.

[10] J. A. Sethian. *Level set methods and fast marching methods*. C. U. P., 1999.

[11] J. Sharpe, U. Ahlgren, P. Perry, B. Hill, A. Ross, J. Hecksher-Sorensen, R. Baldock, and D. Davidson. Optical projection tomography as a tool for 3D microscopy and gene expression studies. *Science*, 296(5567):541–545, Apr 19 2002.

# Fast Segmentation of Focal Liver Lesions in Contrast-Enhanced Ultrasound Data

Spyridon Bakas[1]
S.Bakas@kingston.ac.uk
Katerina Chatzimichail[2]
Katerina@hcsl.com
Bastien Labbé[3]
Bast.Labbe@gmail.com
Gordon J. A. Hunter[1]
G.Hunter@kingston.ac.uk
Paul S. Sidhu[4]
Paulsidhu@nhs.net
Dimitrios Makris[1]
D.Makris@kingston.ac.uk

[1] Digital Imaging Research Centre
Faculty of Science, Engineering & Computing
Kingston University - London - U.K.
[2] Evgenidion Hospital
National & Kapodistrian University
Athens - Greece
[3] Acquisition & Image Processing
Télécom Physique Strasbourg
Strasbourg - France
[4] King's College Hospital
Denmark Hill - London - U.K.

**Abstract**

Assessment of focal liver lesions (FLLs) in Contrast-Enhanced Ultrasound data requires initialisation tasks that are currently performed manually by experienced radiologists. These tasks lead to subjective results, are time-consuming and prone to misinterpretation and human error. This paper describes an attempt to improve this clinical practice by proposing a complete pipeline for the automation of initialisation tasks, such as the identification of a frame where the FLL is well-distinguished, the segmentation of the FLL and the conical area including the ultrasonographic image. The currently proposed novel contribution to automate the FLL segmentation is a fast two-step method, initialised only by a single seed-point, which firstly approximates the FLL by an ellipse and then further refines its shape by iteratively classifying boundary pixels.

## 1 Introduction

Around 47,000 deaths occur annually in the EU due to primary liver cancer, the incidence rate of which is 6.2 and 2.4 per 100,000 people for males and for females, respectively [3]. Contrast-Enhanced Ultrasound (CEUS) is recognised as the most cost-efficient imaging solution for distinguishing between benign and malignant focal liver lesions (FLLs) [11], since it is easy to perform, and uses portable and relatively low cost equipment that allows its presence in every clinic and even at the bedside. Furthermore, CEUS has gained acceptance for the detection and characterisation of very small FLLs –i.e. at the very early/premature stages– with diagnostic accuracy exceeding 95% for the evaluation of malignancies [9].

Radiologists routinely detect, localise, monitor and evaluate FLLs in CEUS video recordings manually, through a very time-consuming series of tasks. These include the identification of a reference frame, where the FLL is sufficiently represented and well-distinguished

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

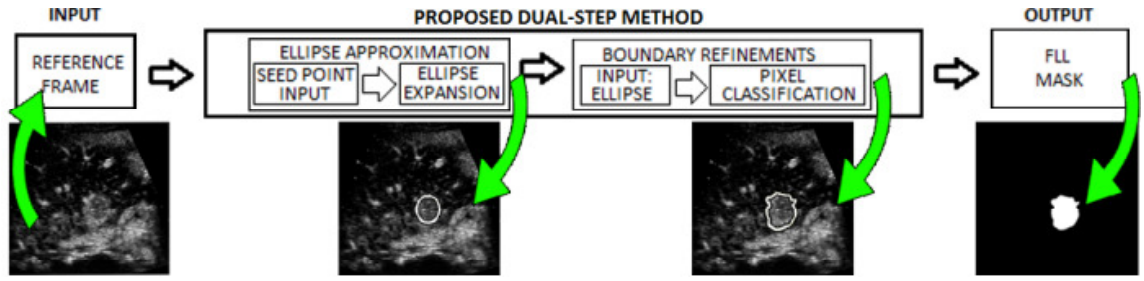



Figure 1: Pipeline of the proposed two-step method for the FLL segmentation.

from the remainder of the image plane in order to manually segment it, and eventually classify it as benign or malignant by monitoring the relative changes of brightness intensity of different regions of interest (ROIs) over time, i.e. their dynamic behaviour [12]. Each of these tasks requires a high level of expertise and provides operator-dependent results, which discourage radiologists from using the technique.

The tasks of monitoring and evaluating the dynamic behaviour of an FLL over the duration of a CEUS recording have been addressed by different solutions [1, 7, 10]. However, all of them assume the existence of prior initialised ROIs, such as the FLL and a larger area (i.e. US mask) including both the FLL and the surrounding healthy liver tissue (i.e. the parenchyma) in a reference frame. The selection of such a frame, as well as the automatic segmentation of the US mask have been addressed in [2] and [1], respectively. However, no solution has ever been suggested specifically for segmenting the FLL shape in a single frame of a CEUS recording. Thus, an expert is still needed to manually annotate the FLL in a frame, after following the standard care protocol. This manual procedure leads to subjective results, is time-consuming, and prone to misinterpretation and human error.

Therefore, automation of the process is highly desirable. Current approaches for segmentation of different tissues in medical images (e.g. snakes [4], level-sets [5]) are impractical due to either slow convergence through the need to optimise a large number of coefficients, or their dependence on a manual initialisation close to the actual boundaries [6]. This paper describes a complete pipeline to automate the initialisation tasks required for evaluating an FLL, whilst focusing on the FLL segmentation, in an attempt to improve the clinical practice by assisting the radiologists to make a diagnosis more easily and with greater confidence.

## 2 Methodology

The proposed pipeline first segments the US mask, by considering an intensity change detection as in [1] and then automatically identifies the optimal reference frame as in [2]. According to radiologists, this frame is expected to be the one with the maximum contrast between FLL and parenchyma.

A novel two-step method is then proposed to segment the FLL boundaries on the reference frame. Firstly, an enhanced version of a fast active ellipse model (AE) is used to approximate the FLL by an ellipse, initialised only by a single seed-point. Such an approximation is appropriate due to the FLL being roughly an ellipsoid by definition [11]. Thus, any 2D representation of an FLL results in an approximately elliptical shape. Secondly, the FLL segmentation is further refined automatically by iteratively classifying boundary pixels according to a probabilistic model of the FLL and the parenchyma intensities.

## 2.1 Ellipse Approximation

A Gaussian distribution model ($G(I_{AE})$) with mean $\mu$ and standard deviation $\sigma$ is used to describe the pixel intensities within the AE. The concept of this AE is to iteratively expand or contract on each major axis independently, according to four different forces.

An AE was first introduced for medical ultrasound in [6], where the forces were based on modified Gaussian functions. However, such smooth functions may not be appropriate to distinguish the FLL boundaries precisely. Furthermore, contraction was not allowed, but only expansion, as its force functions were non-negative (Fig. 2(a)). Therefore, we propose alternative force functions for updating the AE on each iteration based on a rectangular function ($Rect_{2\sigma}(I(p_k))$), which is positive for pixel intensities within one standard deviation $\sigma$ of the mean $\mu$ of $G(I_{AE})$, but negative otherwise.

$$Rect_{2\sigma}(I(p_k)) = \begin{cases} +1 & \text{if } \mu-\sigma \leq I(p_k) \leq \mu+\sigma \\ -1 & \text{if } I(p_k) \leq \mu-\sigma \text{ , or } I(p_k) \geq \mu+\sigma \end{cases} \tag{1}$$

where $I(p_k)$ is the intensity of the pixel $p_k$ and $2\sigma$ specifies the width of the rectangular function. After considering that the boundary of the AE is divided in $K$ equally-spaced points, then the pixel $p_k$ is in the $k^{th}$ of the $K$ points in this boundary.

The sharpness of $Rect_{2\sigma}(I(p_k))$ is more appropriate for emphasising the FLL's boundaries than the modified Gaussian function proposed in [6], and should allow faster expansion and contraction. The force proposed here is defined as:

$$f_g = \sum_{k=1}^{K} Rect_{2\sigma}(I(p_k)) \cdot W_g\left(\frac{k}{K}\right) \quad , \quad \forall g \in \{x_+, x_-, y_+, y_-\} \tag{2}$$

where $x$ and $y$ are the horizontal and vertical forces with the '$+$' and '$-$' subscripts depicting the direction of each of them, and $W_g$ are the weightings used on each force as defined in [6].

The method is initialised by one seed-point within the FLL region. An elliptical shape model $e(x_c, y_c, r_x, r_y)$ is then employed, where $x_c, y_c$ define the centre of gravity (CoG) and $r_x, r_y$ the semi-major axes of the AE. For updating the CoG and the semi-major axes of the AE, at each iteration $i$, we propose the following equation.

$$\begin{bmatrix} \Delta x_{c,i} \\ \Delta y_{c,i} \\ \Delta r_{x,i} \\ \Delta r_{y,i} \end{bmatrix} = \begin{bmatrix} \alpha & -\alpha & 0 & 0 \\ 0 & 0 & \alpha & -\alpha \\ \alpha & \alpha & 0 & 0 \\ 0 & 0 & \alpha & \alpha \end{bmatrix} \begin{bmatrix} f_{x+} \\ f_{x-} \\ f_{y+} \\ f_{y-} \end{bmatrix} \tag{3}$$

where each $\Delta$ represents the change in the corresponding quantity during that iteration and $\alpha$ is an acceleration parameter. This allows an automated update of the AE, based entirely on the forces obtained directly from the intensity distribution of the image.

Updating the position and shape of the AE terminates if either a maximum number of iterations is reached or if the sum of the magnitudes of the forces and the values of $\mu$ and $\sigma$ all converge, i.e. the changes in them each stay below a specified tolerance over three successive iterations. The final output ellipse ($e_{AE}$) of the AE process covers the area $A_{AE}$.

## 2.2 Boundary Refinements using a Probabilistic Model

This probabilistic boundary refinement (PBR) step is initialised by $e_{AE}$ and considers a non-parametric generic shape model of a closed contour with $K$ points. Initially, another ellipse is created ($e_{EXT}$), concentric with $e_{AE}$ but with double its semi-major axes and covering an area $A_{EXT}$ (Fig. 2(b)). The annular-shaped area $A_{ring} = A_{AE}^{C} \cap A_{EXT}$ is then defined, and the pixel populations inside each of $A_{AE}$ and $A_{ring}$ are assumed to describe the classes of the FLL and parenchyma, respectively, i.e. $C_F$ or $C_P$. The segmentation of the FLL's boundary is then treated as a binary classification problem, allocating pixels to either $C_F$ or $C_P$.

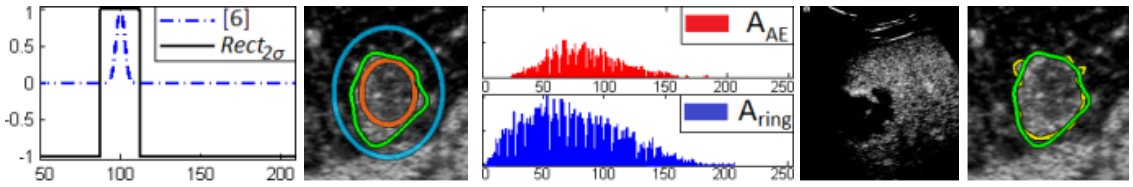



Figure 2: (a) shows the 2 different force functions applied to the AE. In (b) the red (inner) ellipse is $e_{AE}$ and the concentric blue (outer) ellipse is $e_{EXT}$. The green (irregular-shaped) contour depicts the final segmentation of the FLL by the proposed method ($F_d$). (c) shows the histogram of the pixel brightness intensities of the two areas of interest ($A_{AE}$ and $A_{ring}$). (d) shows an FLL not well-distinguished due to acoustic shadow. (e) gives the comparison of $F_d$ (green) compared with the manually annotated ground truth $F_{GT}$ (yellow) giving $J = 84\%$.

Two distinct Gaussian probability distributions are used to model the pixel intensities inside $A_{AE}$ and inside $A_{ring}$, respectively (Fig. 2(c)). Specifically, the proposed approach processes samples from the population of each class separately, and models the class-conditional probabilities ($Pr(I(p_k)|C_F)$ and $Pr(I(p_k)|C_P)$) by fitting a Gaussian model, using Maximum Likelihood Estimation to find the mean and standard deviation, for each class.

Consequently, at each iteration, the intensity $I(p_k)$ of each boundary point $p_k$ is assessed and $Pr(I(p_k)|C_F)$ compared with $Pr(I(p_k)|C_P)$. This point $p_k$ is classified into $C_F$ or $C_P$ according to the larger of the conditional probabilities. (This is equivalent to a Bayesian classification with equal priors for the two classes.) The contour is then expanded or contracted along the radial line from the CoG to $p_k$, based on the above classification decision.

The updating of the position and shape of the closed contour terminates if either a maximum number of iterations is reached or if the contour's length converges over five successive iterations, i.e. the changes in the length stay below a specified tolerance.

# 3 Data and Evaluation Metrics

Data of 60 real clinical cases of a multi-centre study with patients in different physical conditions were provided. 5 cases had to be excluded from the evaluation since the FLL was not well-distinguished from the remainder of the image plane due to acoustic shadows (Fig. 2(d)). All data was acquired using Siemens ACUSON Ultrasound (US) systems (Mountainview CA). 46 cases were captured at King's College Hospital in the UK, at spatial resolution $1024 \times 768$ pixels, using an S2000 US system equipped with 4 (or) 6 MHz curvilinear transducer. The remaining 14 cases were captured at Evgenidion Hospital in Greece, at spatial resolution $768 \times 576$ pixels, using a Sequoia C512 US system equipped with 6-2 MHz curvilinear transducer. In all examinations the second generation contrast medium SonoVue [8] (Bracco S.p.A., Italy) was used in a 2.4ml bolus intravenous injection (into an arm vein), which allows excellent depiction of the FLL vascularity and perfusion [12]. Specific acquisition parameters of the equipment for each patient are unknown, as they were set by the radiologist individually at the start of each examination. The acquisition method of this data reflects true clinical practice and leads to increased variability. Examinations were performed by radiologists with 13-16 years of experience using CEUS. All data were obtained without prior knowledge of subsequent processing by a software tool and without any specific instructions being given to the radiologist beforehand. Appropriate ethics and confidentiality procedures have been followed at all times.

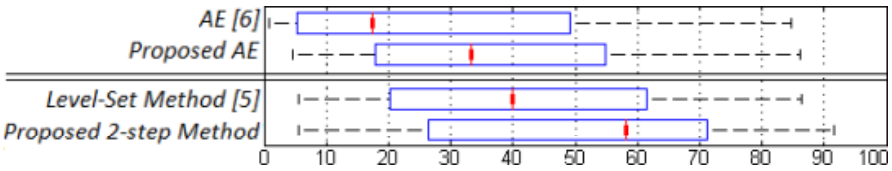



Figure 3: Jaccard index for assessing the overlap between $F_{GT}$ and $F_d$ for the original AE [6], the Level-Set method [5] and the proposed method across our datasets.

To evaluate quantitatively the proposed method at the pixel level, the ground truth silhouettes of the FLLs ($F_{GT}$) were manually annotated by a radiologist in the reference frame. These silhouettes are then compared with the areas segmented by the proposed method, $F_d$ (Fig. 2(e)). The metric considered for evaluating the similarity and difference between $F_{GT}$ and $F_d$ is the Jaccard index ($J = \frac{|F_{GT} \cap F_d|}{|F_{GT} \cup F_d|}$), as it strictly accepts information only from pixels $p_{x,y} \in F_{GT} \cap F_d$ and penalises pixels misclassified as either FLL or non-FLL ($p_{x,y} \in F_{GT} \triangle F_d$).

# 4 Experiments and Results

Our proposed methods, along with the original AE [6] and a Level-Set method [5] for comparison, were implemented in Matlab v.7.12 without parallelism on an Intel i7-2620M platform with 8GB RAM and applied to 55 cases where the FLL was well-distinguished. The value of the acceleration parameter $\alpha$ in Eq.3 was set equal to 1 in all cases.

Comparing the median values and inter-quartile ranges (IQR) of $J$ between our enhanced AE and the AE of [6] justifies the proposed improvements (Fig. 3). In addition to the principal improvement, which is the increase in the value of $J$, the enhanced AE needs on average only 59.4% of the time required for the AE of [6] to converge, namely 1.2 rather than 2 seconds on average per case, as the rectangular force functions lead to faster ellipse fitting. This could be significant where a large number of cases are processed.

Applying the proposed PBR step to the output from the AE ($e_{AE}$) gives further improvement to $J$, indicating a much better segmentation of the FLL (Fig. 2(b)). The results from the application of PBR are compared with results of a Level-Set method [5], again initialised by $e_{AE}$. Both the median and IQR of $J$ across all cases are higher for PBR than for the Level-Set method, showing that the FLL segmentation is superior when using the former. Last but not least, our method converges in a small fraction (2.2%) of the time required by the Level-Set method, e.g. 5.9 seconds rather than 4.4 minutes on average per case.

# 5 Conclusions and Future Work

In this paper, we proposed a complete pipeline to automate the initialisation tasks required for the evaluation of FLLs in CEUS data. In particular, a novel two-step method was proposed for automating the FLL segmentation based on a probabilistic model. The first step consists of an enhanced AE using a rectangular force function, to approximate the FLL by an ellipse. The second step is used for refining this approximation by iteratively classifying boundary pixels according to a probabilistic model. Experimental comparison of the proposed version of AE to the original one [6] justifies the proposed enhancements. In addition, the second step of our method demonstrated better performance and lower computational time when compared with Level-Sets [5]. However, its performance is dependent on the manual input of one seed-point and the ROI should be in a form which can be reasonably well-bounded by a closed contour. The proposed method should also be appropriate to other modalities and application areas, e.g. breast US. Last but not least, coupling of the proposed pipeline

with an existing tracking method [1, 7, 10] could lead to an automated framework for FLL evaluation providing a second-opinion tool to radiologists.

# References

[1] S. Bakas, A. Hoppe, K. Chatzimichail, V. Galariotis, G. Hunter, and D. Makris. Focal Liver Lesion Tracking in CEUS for Characterisation Based on Dynamic Behaviour. *Springer, Advances in Visual Computing*, LNCS(7431):32–41, 2012.

[2] S. Bakas, G. Hunter, C. Thiebaud, and D. Makris. Spot the Best Frame: Towards Intelligent Automated Selection of the Optimal Frame for Initialisation of Focal Liver Lesion Candidates in Contrast-Enhanced Ultrasound Video Sequences. *9th International Conference on Intelligent Environments (IE'13 - Athens, Greece)*, pages 196–203, 2013.

[3] M. Blachier, H. Leleu, M. Peck-Radosavljevic, D.-C. Valla, and F. Roudot-Thoraval. The Burden of Liver Disease in Europe - A Review of Available Epidemiological Data. *European Association for the Study of the Liver*, 2013.

[4] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[5] C. Li, C. Xu, C. Gui, and M. D. Fox. Distance Regularized Level Set Evolution and Its Application to Image Segmentation. *IEEE Trans Image Proces*, 19(12):3243–54, 2010.

[6] M. Marsousi, J. Alirezaie, A. Ahmadian, and A. Kocharian. Segmenting Echocardiography Images using B-Spline Snake and Active Ellipse Model. *Conf Proc IEEE Eng Med Biol Soc*, 2010:3125–3128, 2010.

[7] N. Rognin, R. Campos, J.-P. Thiran, T. Messager, A. Broillet, and P. Frinking et al. A New Approach For Automatic Motion Compensation For Improved Estimation of Perfusion Quantification Parameters in Ultrasound Imaging. *The 8th French Conference on Acoustics (Tours, France)*, pages 61–65, 2006.

[8] M. Schneider. Charasteristics of SonoVue. *Echocardiography*, 16(7):743–746, 1999.

[9] D. Strobel, K. Seitz, W. Blank, A. Schuler, C. F. Dietrich, and A. vonHerbay et al. Tumor-Specific Vascularization Pattern of Liver Metastasis, Hepatocellular Carcinoma, Hemangioma and Focal Nodular Hyperplasia in the Differential Diagnosis of 1349 Liver Lesions in Contrast-Enhanced Ultrasound. *Ultraschall Med*, 30(4):376–82, 2009.

[10] C. N. Ta, Y. Kono, C. V. Barback, R. F. Mattrey, and A. C. Kummel. Automating Tumor Classification With Pixel-by-Pixel Contrast-Enhanced Ultrasound Perfusion Kinetics. *J Vac Sci Technol B Nanotechnol Microelectron*, 30(2):02C103–01–10, 2012.

[11] M. E. Westwood, M. A. Joore, J. P. C. Grutters, W. K. Redekop, N. Armstrong, and K. Lee et al. Contrast-Enhanced Ultrasound Using SonoVue®(Sulphur Hexafluoride Microbubbles) Compared With Contrast-Enhanced Computed Tomography and Contrast-Enhanced Magnetic Resonance Imaging for the Characterisation of Focal Liver Lesions and Detection of Liver Metastases: a Systematic Review and Cost-Effectiveness Analysis. *Health Technol Assess*, 17(16), 2013.

[12] S. R. Wilson and P. N. Burns. Microbubble-Enhanced US in Body Imaging: What Role? *Radiology*, 257(1):24–39, 2010.

# A Novel Breast Image Preprocessing For Full Field Digital Mammographic Segmentation and Risk Classification

Wenda He[1]
weh@aber.ac.uk
Minnie Kibiro[2]
minnie.kibiro@nnuh.nhs.uk
Arne Juette[2]
arne.juette@nnuh.nhs.uk
Peter Hogg[3]
p.hogg@salford.ac.uk
Erika R. E. Denton[2]
erika.denton@nnuh.nhs.uk
Reyer Zwiggelaar[1]
rrz@aber.ac.uk

[1] Department of Computer Science, Aberystwyth University, Aberystwyth, SY23 3DB, UK
[2] Department of Radiology, Norfolk & Norwich University Hospital, Norwich, NR4 7UY, UK
[3] School of Health Sciences, University of Salford, Salford, M6 6PU, UK

**Abstract**

To obtain optimal breast image quality during the image acquisition, a compression paddle is used to even the breast thickness. Clinical observation has indicated that breast peripheral areas may not be fully compressed, and may cause unexpected intensity and texture variation within these areas. Such breast parenchymal appearance discrepancies may not be desirable for tissue modelling within computer aided mammography. This paper describes a novel mammographic image preprocessing method to improve the image quality before analysis. Mammographic segmentation and risk classification were performed to facilitate a quantitative and qualitative evaluation, using digital mammographic images. Visual assessment indicated significant improvement on segmented anatomical structures and tissue specific areas when using the processed images. The achieved risk classification accuracies are 80% and 79% for Birads and Tabár risk scheme, respectively. The developed method has demonstrated an ability to improve the quality of mammographic segmentation, leading to more accurate risk classification. This in turn can be found useful in early breast cancer detection, risk-stratified screening, and aiding radiologists in the process of decision making prior to surgery and/or treatment.

## 1 Introduction

Within screening mammography, a compression paddle is used to even out the breast tissue in order to obtain high quality mammographic images. However, when the breast is subjected to compression, the peripheral areas may not be compressed due to a reduction

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

of breast thickness. This results in air gaps above and beneath the uncompressed areas, leading to X-ray scattering, degradation in contrast, and limitation of the quantitative usefulness of radiographic images [2]. Compression paddle related image quality issues vary due to difference in breast size and composition; *e.g*. in extreme cases, the visibility of breast parenchyma is too low in the peripheral areas to be examined. Therefore, it is necessary to develop an image processing method to improve the visibility of peripheral uncompressed area of the projected breast, which facilities presentation and can be beneficial to mammographic analysis [1]. Image processing (enhancement) methods developed to address the aforementioned problem can be categorised into two groups; parametric [1, 8] and non-parametric [9] approaches. Methods [8] and [9] are critically dependent on the accuracy of interactive segmentation of the dense breast tissue and fatty tissue interpolation, whilst [9] can only be applied to unprocessed digital mammograms before logarithmic transformation. The reader is referred to [1, 8, 9] for the details of the developed methodologies.

Regarding mammoraphic risk assessment, Tabár *et al*. [10] proposed a model based on a mixture of four mammographic building blocks representing normal breast anatomy, mammographic risk is categorised into five classes along these building blocks (*i.e*. [nodular%, linear%, homogeneous%, radiolucent%]). Pattern I [25%, 15%, 35%, 25%]; pattern II/III [2%, 14%, 2%, 82%]; pattern IV [49%, 19%, 15%, 17%]; and pattern V [2%, 2%, 89%, 7%] [10]. Alternatively, American College of Radiology's Breast Imaging Reporting and Data System (Birads) [6] was developed, mammographic risk is categorised based on the percentage of dense breast tissue in four risk classes: Birads 1, the breast is almost entirely fat ($< 25\%$ glandular); Birads 2, the breast has scattered fibroglandular densities ($25\% - 50\%$); Birads 3, the breast consists of heterogeneously dense breast tissue ($51\% - 75\%$); and Birads 4, the breast is extremely dense ($> 75\%$ glandular).

We present a novel mammographic image preprocessing technique, which models a breast by estimating relative breast thickness ratios using both Mediolateral Oblique (MLO) and Cranio-Caudal (CC) views, the correction process is marginally similar to [5, 11]. An additional automatic selection method was developed to better target images requiring enhancement in a systematic way. A quantitative and qualitative evaluation was conducted to assess the robustness of the method; all processed images were subjected to mammographic segmentation and subsequent risk classification using both Tabár and Birads risk scheme.

## 2 Data and Method

A total of 360 digital mammographic images were used in the experiment. A consensus ground truth was obtained based on three radiologists. Modelling Tabár's mammographic building blocks, a collection of mammographic patches consisting of a total of 344, 89 and 457 examples of nodular, homogeneous and radiolucent (similar to Birads dense, semi-dense and non-dense) tissue was cropped to various sizes from randomly selected images from the dataset. Note that the linear structure was not specifically included in the experiment, instead this tissue type was considered part of the other tissue classes as it appears in combination with all three classes. Therefore, only nodular, homogeneous and radiolucent tissues were modelled and used in the segmentation process.

The developed methodology starts off with automatic image selection, the image preprocessing consists of four steps; 1) X-ray penetration probability weighting map generation, 2) intensity balancing, 3) intensity ratio propagation, and 4) boundary stitching. All the processed images were subjected to geometric moments based mammographic segmentation [3]

and risk classification.

## 2.1 Automatic Image Selection

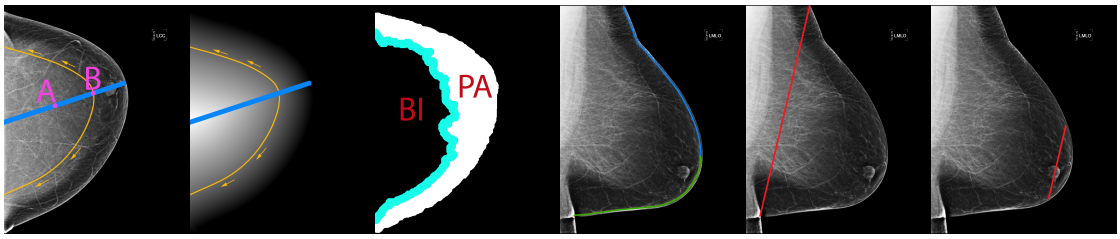
Thorough empirical observations of the collected data indicated that a mammographic image may require processing if its automatically calibrated parameters *acp* (*i.e.* compression force (*CF*), breast thickness (*BT*) and peak kilovotage (*KVP*)) are less than the mean values. It was less robust to use the *acp* alone as a constraint due to large breast tissue density, composition and size variations. Additional constraints based on prior knowledge were incorporated to strengthen the selection robustness. Specifically, Otsu [7] automatic binary image segmentation technique was used to separate the peripheral area (*PA*) and breast interior (*BI*). Three properties were calculated: total peripheral area (*TPA*), vertical peripheral area coverage (*VPAC*) in image rows and pectoral coverage (*PC*) in image rows. A mammographic image (*img*) was selected to be processed if $CF_{img} <= \overline{CF}$ & $BT_{img} <= \overline{BT}$ & $KVP_{img} <= \overline{KVP}$ & $15\% < TPA_{img} < 50\%$ & $VPAC_{img} > 75\%$ & $PC_{img} < 30\%$. The threshold values were empirically defined through trial and error, to achieve the best separation overall.

## 2.2 Mammographic Image Preprocessing

The X-ray penetration probability has a direct correlation with breast thickness. Due to physical complexity (*e.g.* unknown combination factors in X-ray beam spectrum and breast tissue composition), the X-ray penetration probability was modelled in a simplified way by encompassing all other elements (*e.g.* dosage, filter, anode) in a "black box" only considered as inversely proportional to the breast thickness. An X-ray penetration probability weighting map was generated for each image by calculating and propagating the relative breast thickness ratios based on its pair. For example, to a CC view, the relative breast thickness ratio (*r*) can be estimated based on the projected physical contour of the compressed breast as seen on MLO view. In particular, the skinline was firstly extracted from the MLO view and split in two at the furthest pixel to the chest wall to form upper and lower skinlines (*e.g.* blue and green lines in Figure 1 (c)). For each pixel in the top skinline, a corresponding pixel was sought in the lower skinline, to form a parallel line (*p-line*) (*e.g.* red line in Figure 1 (d)) to the chest wall by linking the two pixels. This process was repeated for all the pixels in the top skinline, and resulting in a series of parallel lines (*e.g.* Figure 1 (d)(e)) . To the CC view, the *r* at a given point (*p*) (*e.g.* 'A' in Figure 1 (a)) is calculated based on the boundary pixel ($p_{base}$) (*e.g.* 'B' in Figure 1 (a)) which separates *PA* and *BI* as $r = \frac{p-line(p)}{p-line(p_{base})}$); both pixels are on the thickest projected section (*e.g.* blue lines in Figure 1 (a)(b)) on the CC view. To complete the X-ray penetration probability weighting map for the CC view, the remaining pixels on the thickest projected section with the estimated breast thickness ratios were assigned in the same way, and the calculated ratios were propagated to the pixels that have the same distance to the skinline (S) (*e.g.* pixels on the yellow lines in Figure 1 (a)(b)) .

To reduce the intensity distribution variation, a base weight ($w_{base}$) was firstly calculated as $w_{base} = \frac{\sum_{i=0}^{N} W_i(x,y)}{N}$, $\forall W_i(x,y) \in S$, where *W* and *S* denote the weighting map and the boundary between *PA* (uncompressed breast peripheral area) and *BI* (breast area not part of *PA*). For each pixel within the *BI*, the intensity value *P(x, y)* was altered as $P'(x,y) = \frac{w_{base}}{W(x,y)} P(x,y)$.

After the intensity balancing, the local intensity ratio was propagated as a means of improving tissue appearance in *PA* (similar to the process described in [11]). From pixels at the boundary *S* to the skinline within the *PA*, each intensity value was altered by calculating




(a) (b) (c) (d) (e) (f)

Figure 1: Image illustration, from left to right: (a) CC view, (b) its distance map (pixel to skinline), (c) areas of *PA*, *BI*, and *S* (cyan line), (d) pair MLO view, (e) parallel lines proximal to pectoral muscle and (f) proximal to nipple.

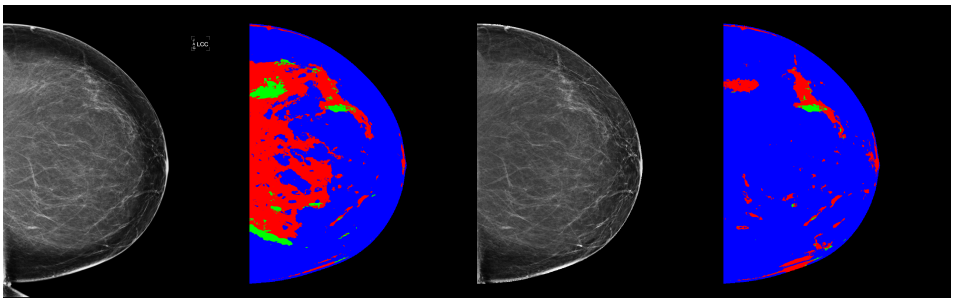
the propagation ratio (*pr*) for each pixel *P(x, y)* with distance to the skinline *D(x, y)* and within an empirically defined 17 $\times$ 17 neighbourhood for the efficiency and robustness, as $I_{avgP_1} = \frac{\sum_{j=0}^{M} P_j(x,y)}{M} \forall P_j(x,y) = D(x,y)+1$, $I_{avgP_2} = \frac{\sum_{i=0}^{N} P_i(x,y)}{N} \forall P_i(x,y) = D(x,y)+2$, $pr = \frac{I_{avgP_2}}{I_{avgP_1}}$, and $P'(x,y) = pr \times P(x,y)$; where $D(x,y)+1$ and $D(x,y)+2$ are pixel distances to skinline 1 and 2 steps further away from the observed pixel.

Boundary stitching was applied to seamlessly normalise pixel intensity within boundary *S*, thickened to 5 pixels band, in order to gradually smooth the transition from *BI* to *PA*; the maximum and minimum intensity values were determined within an empirically defined 7 $\times$ 7 neighbourhood (see Section 3 for the effects of using different neighbourhood sizes).

## 2.3 Mammographic Segmentation and Risk Classification

A geometric moments based mammographic segmentation [3] was applied to all the images. To be concise, this method is used to extract texture features from a set of mammographic patches using geometric moments; the derived feature vectors are expected to contain not only texture primitives but also geometric information. The methodology was modified by incorporating a feature and classifier selection process using a collection of attribute selection algorithms and classifiers available in Weka [4] through four stages: 1) A set of neighbourhoods (*i.e.* {7, 17, 27, 37, 47, 57, 67}) covering small to large anatomical structures were predefined and used in the feature extraction process over mammographic patches containing different breast tissue examples (*e.g.* nodular, homogeneous and radiolucent). 2) The raw feature vectors for all the patch pixels were filtered utilising all available filtering methods in Weka for reduction of attributes. Empirical testing indicated that the most frequent and prominent features are within the top half of the output attributes list after filtering. 3) The filtered feature vectors were then subjected to all available attribute selection (wrapper) methods in Weka to select the most discriminative subset of the filtered features. The most frequent attributes in the output subsets were used as the final selected optimal features. 4) All the available classifiers in Weka were used to perform (10-fold) cross-validation based evaluation over the selected optimal features. The classifier achieving the highest classification results (random committee in this study) was used in conjunction with the selected features in the mammographic segmentation using Tabár tissue modelling.

An unseen pixel's tissue class is determined by the trained classifier. The relative proportions of the mammograpic building blocks (tissue composition) were calculated from the resultant mammographic segmentation as features, and used in a leave one-image out (10-fold) cross-validation as a means of risk classification.



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

Figure 2: (a) original image (Tabár II/Birads 1), (b) segmentation of (a) with over segmented glandular (*i.e.* red nodular and green homogeneous) tissue, (c) processed image, (d) segmentation of (c) showing more segmented blood vessels in the peripheral area, segmented glandular tissue is more realistic and fatty tissue (blue) is more correctly identified than (b).

| Birads | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 149 | 21 | 0 | 0 |
| 2 | 17 | 100 | 6 | 0 |
| 3 | 0 | 12 | 23 | 8 |
| 4 | 0 | 0 | 7 | 17 |

| Tabár | I | II/III | IV | V |
|---|---|---|---|---|
| I | 122 | 21 | 4 | 0 |
| II/III | 19 | 128 | 1 | 0 |
| IV | 13 | 0 | 24 | 8 |
| V | 1 | 0 | 8 | 11 |

Table 1: Risk classification confusion matrices; Birads left, Tabár right.

## 3 Results and Discussion

Visual assessment was conducted to assess the quality of enhanced images. The majority of cases showed processed images to have improvements in textural appearance and contrast in the peripheral areas as expected; see Figure 2 (a)(c) for example. However, over and under enhancement may occur in some cases if *PA* and *BI* are not separated correctly, or the breast thickness ratios are wrongly estimated. During boundary stitching, undesirable artefacts may be created when a larger neighbourhood is used. Incorrect local intensity alternation affects textural appearance which can be perceived as an artefact. However, the processed images seem to have minimum textural distortion and were suitable for the follow up image analysis. The segmentation results have shown significant improvements in terms of correctness of the segmented anatomical structures over the breast parenchyma and tissue specific areas.

The risk classification accuracies increased on average 7% when compared with the results obtained before the image preprocessing was applied. The total classification accuracies were 80% and 79% for Birads and Tabár risk scheme, respectively. Table 1 shows the confusion matrices with respect to the two risk schemes when the developed methodology was applied. In both cases, mammographic images in high risk classes seem to have more misclassification (percentage wise), which may relate to the intensity over balancing for homogeneous tissue with structureless densities. The achieved risk classification results are in line with results achieved by the state-of-art method [11], but different segmentation principle (*e.g.* two class fatty/dense segmentation) and datasets were used. Further segmentation improvement can be made by widening the tissue patch variation, leading to a more robust classifier which in turn could also improve risk classification accuracies. However, this is outside the scope of this study and is considered focus for future investigation on mammographic segmentation and risk classification methodology.

# 4 Conclusions

The developed mammographic preprocessing can be used to reduce intensity and textural appearance discrepancies. Results have shown more anatomically accurate and consistent segmentation over the breast parenchyma when using the processed images in conjunction with the selected feature and classifier. This in turn improved subsequent risk classification accuracies. Incorporating the novel image preprocessing into mammographic segmentation methodology could prove useful in quantifying change in relative proportion of breast tissue, aiding radiologists' estimation in mammographic risk assessment, and providing risk-stratified screening to patients.

# References

[1] J. W. Byng, J. P. Critten, and M. J. Yaffe. Thickness-equalization processing for mammographic images. *Radiology*, 203(2):564–568, 1997.

[2] J. L. Ducote and S. Molloi. Scatter correction in digital mammography based on image deconvolution. *Physics in Medicine and Biology*, 55:1295–1309, 2010.

[3] W. He, E. R. E. Denton, K. Stafford, and R. Zwiggelaar. Mammographic image segmentation and risk classification based on mammographic parenchymal patterns and geometric moments. *Biomedical Signal Processing and Control*, 6(3):321–329, 2011.

[4] E. Frank I. H. Witten and M. A. Hall. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, 3 edition, 2011.

[5] N. Karssemeijer and G. Te Brake. Combining single view features and asymmetry for detection of mass lesions. *Computational Imaging and Vision*, 13:95–102, 1998.

[6] American College of Radiology. *Breast Imaging Reporting and Data System BI-RADS*. Reston, VA: American College of Radiology, 4th edition, 2004.

[7] M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13:146–165, 2004.

[8] P. R. Snoeren and N. Karssemeijer. Thickness correction of mammographic images by means of a global parameter model of the compressed breast. *IEEE Transactions on Medical Imaging*, 23(7):799–806, 2004.

[9] P. R. Snoeren and N. Karssemeijer. Thickness correction of mammographic images by anisotropic filtering and interpolation of dense tissue. *SPIE*, 5747:1521–1527, 2005.

[10] L. Tabár, T. Tot, and P. B. Dean. *Breast Cancer: The Art And Science Of Early Detection With Mamography: Perception, Interpretation, Histopatholigic Correlation*. Georg Thieme Verlag, 1 edition, December 2004.

[11] M. Tortajada, A. Oliver, R. Martí, M. Vilagran, S. Ganau, L. Tortajada, M. Sentís, and J. Freixenet. Adapting breast density classification from digitized to full-field digital mammograms. *Lecture Notes in Computer Science*, 7361:561–568, 2012.

# Automatic Reflection Detection on Dermatological Images Acquired via Mobile Devices

Maria João M. Vasconcelos[1]
maria.vasconcelos@fraunhofer.pt
Luís Rosado[1]
luis.rosado@fraunhofer.pt

[1] Fraunhofer Portugal AICOS
Porto, Portugal

**Abstract**

The late detection of skin cancer, especially melanoma, leads to an increase in mortality rates, therefore it is high time to embrace mobile teledermatology and adapt the existing methodologies to check out the malignancy level of a mole. Though, the usage of a smartphone camera gives rise to new pre-processing challenges like the treatment of reflections caused by the flash light or uneven illumination conditions. In this work, a new methodology to detect reflections on mobile-acquired dermatological images is presented, based on the different behaviour in terms of highlights invariance of the L and H colour channels, from the CIE L**a***b** and HSV color spaces, respectively. The achieved results confirm the good quality of the suggested algorithm.

# 1 Introduction

With increasing aging population globally, there is a growing incidence of skin cancer. Nevertheless the percentage of population participating in skin cancer screening versus incidence is alarmingly low. Late detection leads to rise in skin cancer mortalities, especially melanomas, as so there is a need for complementing existing technologies in order to check out the malignancy level of a mole. In addition, given the current need to decrease the costs of the healthcare providers, the usage of new lightweight monitoring systems which can be carried around easily and used regularly by the patients, is considered crucial to find new ways of making better decisions on treatment.

Considering the high dissemination of mobile devices and their integrated image acquisition systems, mobile teledermatology appears as a promising tool for personal dermatologic data acquisition. When compared to dermoscopy [1], the current most accepted method for image acquisition in dermatology, the usage of a smartphone camera brings up a new set of challenges. Consequently, it is important to develop methodologies to deal with some of the expected artifacts like reflections, blurred images, and uneven illumination conditions. The present work focuses on the detection of reflections in dermatological images acquired via mobile devices.

The most known and commonly used color space in computer applications is the RGB (Red-Green-Blue) which are three primary additive colors and are represented by a three-

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

dimensional, Cartesian coordinate system. In an attempt to linearize the perceptibility of color differences, the CIE proposed two other color spaces, CIE L*u*$v$* and CIE L*a*b*. The HSL (hue, saturation and lightness) color space or similar ones like HSI (intensity), HSV (value) or HCI (chroma) are linear transformations from RGB. The main disadvantage of the RGB color space in applications with natural images is the high correlation between its components. The advantage of CIE absolute color spaces is that they define color exactly and thus includes more color than other color spaces, even more than the human eye can see, however they suffer from being quite unintuitive. In summary, different color spaces are better for different applications, since some colors are perceptually linear or just more intuitive to use, also some color spaces are tied to a specific equipment while other are equally valid on whatever device they are used [2].

Many different approaches to detect reflections have also been proposed over the years. One of the most popular model is the dichromatic reflection model proposed by Shafer 1985 [3], it separates reflectance into surface body reflectance and interface reflectance. However the application of this model to uncalibrated real-word images led to multiple problems. In [4] an overview of the methods to separate reflection components, also known as specularity or highlights removal is given. In [5] the authors use a simple reflection detection algorithm that applies thresholds to the intensity and the difference between the intensity and the average intensity; however it is applied only to dermoscopy images. In [6], the authors introduces a low-level feature for recognizing reflective objects from natural images, also showing how to characterize static specular flows and identify specular surfaces with minimal requirements about the illumination conditions and utilizing a single monocular image.

Recently, [7] presented an automated technique for detecting reflections in image sequences through the analysis of motion trajectories of feature points. In [8] an algorithm for reflections removal specifically for images acquired from mobile phones is presented, however it requires the user to take two pictures, one with flash light and another without in order to produce the new picture with the reflections removed.

The structure of this paper is as follows: the next section describes the dataset used in this work; section 3 details the methodology to detect reflections for dermatological images; section 4 presents and discusses the obtained results; to end, conclusions and future work are summarized.

# 2 Dataset

The database used in our study was obtained under the scope of the Melanoma Detection, a project that intends to create a prototype for a patient-oriented system of skin lesion analysis using a smartphone [9]. The database contains a total of 90 images, 8-bit color with 652x652 pixels of resolution acquired with a mobile phone Htc One S.

Each image of this dataset was later labelled by 5 subjects regarding its reflection level, using a scale from 0 to 3. Subjects were asked to label each image taking only into consideration the mole area, being supplied a complete image and another only with the segmented mole. The final dataset was classified according the most voted class and the images with less than 3 equal votes were discarded from the study. A final dataset of 75 images were used to validate our algorithm, where 16 images were classified with no reflections (class 0), 27 with few reflections (class 1), 22 with some reflections (class 2) and 10 with many reflections (class 3).

# 3 Reflection Detection Algorithm

The first step of this algorithm consists in applying a retina Parvo filtering [10], where the mean luminance is attenuated, the spectrum is whitened and the contours enhanced. The retina model takes into account the processing occurring at the retina level, being developed to prepare images for high level processing and designed to be used in contexts where contours and contrasts constitute important information, which is our case.

Afterwards, the filtered RGB image is transformed into Lab color space to obtain L channel according to equations (1) and (2):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \tag{1}$$

$$L = \begin{cases} 116 * Y^{1/3} - 16 & for \quad Y > 0.008856 \\ 903.3 * Y & for \quad Y \leq 0.008856 \end{cases}. \tag{2}$$

As the studies indicate [2], the H channel is invariant to highlights so we inferred if this condition was still preserved for dermatological images. The previous statement was confirmed in our study nonetheless, after some further exploration on the H channel original equation, our research indicated that the inclusion of slight variation on it brought significant better results concerning highlights invariance. Therefore, the H* channel is calculated from the HLS space taking into consideration the subsequent equation (3):

$$H^* = 60(G - B)/(\max(R, G, B) - \min(R, G, B)) \tag{3}$$

Because $0 \leq H^* \leq 360$, the channel is later normalized to fit between 0 and 1 and be comparable to the L channel. Taking into account that the L channel is considered dependent on highlights, while the H channel is invariant to highlights [2], the difference between the L and H* channel is obtained with the purpose of investigate if it enhanced the reflection regions. And in fact, the results attained with the difference mask were better than the ones using only the H* channel, so we decided to choose for the first option.

In order to remove small artifacts on the differences mask, the image is then smoothed using a Gaussian filter. Median filter and average filter were also tested to smooth the differences mask and turned to be less effective. This can be explained by the fact that neighbor pixels in Gaussian filter have different coefficients related with Gaussian distribution, having the current pixel a more significant contribution. Thus the edges are more preserved in Gaussian filtering, which in our case are the reflections.

Finally, to segment reflection regions, the histogram of the difference mask is computed and the color level with maximum frequency is chosen as the binarization threshold. This threshold is appropriate independently of the fact that the maximum frequency represents the background or the skin mole groups, because the color variation of the background and skin mole pixel groups is lower after Gaussian filtering, and also only reflections inside the mole segmentation mask are considered.

Figure 1 depicts the outputs along the several steps of our algorithm: first we present the original image, followed by the image after the application of the retina filter, then images of the L and H channels are portrait as well as the resulting difference, after that the segmentation of the mole is shown (in white) and the mask of reflections (in black), at last we represent the final mask.

The algorithm was developed using C++ and OpenCV library [11]. To quantify the presence of reflection regions in each image, the relative area of the obtained reflection mask was calculated, by dividing the area of the reflection zones by the total area of each mole. The total area of the skin mole corresponds to the area of the correspondent segmentation mask, which was obtained using a simple adaptive segmentation method [12]. In addition, the number of blobs on each reflection mask was also calculated.

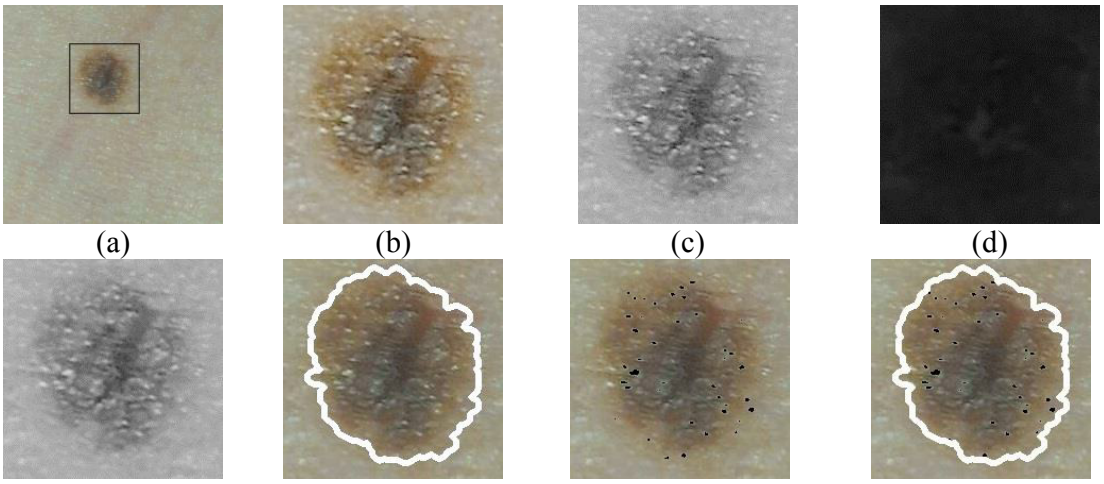
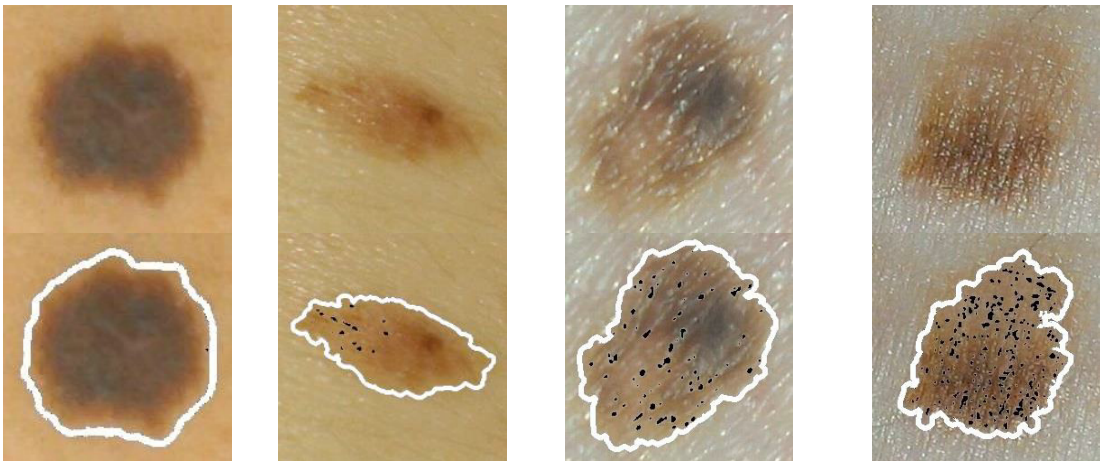



Figure 1: Example of results obtained during the application of the reflection algorithm: (a) original image; (b) after using the retina filter; (c and d) L and H* channels; (e) difference mask; (f) segmentation mask in white; (g) reflection mask in black; (h) final mask.

# 4 Results

In this section we present the results of the proposed algorithm, using the previously referred image dataset. As it was previously said, the images on the used dataset were classified according to its level of reflections. Figure 2 portrays examples of images of each category in the first row and the second row the reflection mask result is presented. From that, the one could confirm that the suggested algorithm shows a promising discriminatory ability for reflection regions; images classified with no or few reflections present reflection masks with a significant lower occurrence of reflection regions, when compared with images classified as having some or many reflections. To ensure the quality of the presented algorithm, besides the visual inspection, the authors calculated the relative area of the obtained reflection mask in addition to the number of blobs. Figure 3 depicts the distribution graphs for each class of images that once again confirm our methodology. Clearly, as the classes advance in terms of the incidence of reflections, the values of relative area and number of blobs arise.

However for class 1 a high maximum was obtained for the relative area and for class 2 the same occurs for the maximum number of blobs. So, we have decided to look the corresponding images in more detail, which can be observed in figure 4. The first image was classified has having few reflections, in fact 3 subjects classified it as having few reflections and the other 2 subjects classified it as class 2, by looking at the original image it is fair to consider that this image was misclassified. The same happens with the other outlier of class 2, the image was classified as class 2 with 3 votes, and the other subjects voted class 3, this image should be considered as class 3.



No reflections (0) Few reflections (1) Some reflections (2) Many reflections (3)

Figure 2: Examples of original images (first row) and resulting reflection detection masks (second row) for each class.

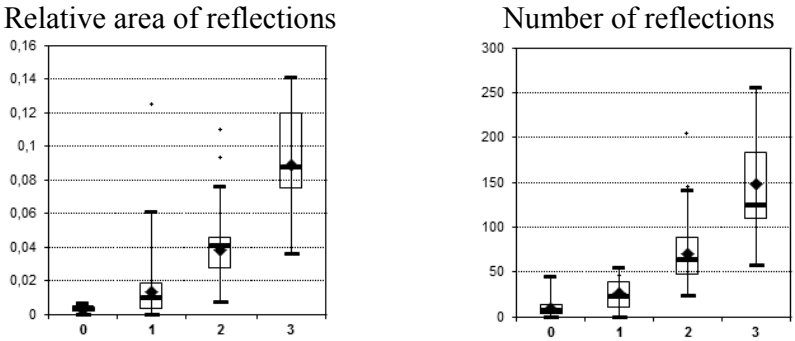



Figure 3: Data distribution for relative area and number of reflections (vertical axis) for each class (horizontal axis).

Figure 4: Images and resulting reflection detection mask of two outliers identified from the data distribution charts.

# 5 Conclusions and Future Work

The development of algorithms for pre-processing analysis especially designed for application in dermatological images acquired from mobile devices is definitely a requisite nowadays. Algorithms capable of doing image quality assessment for these specific images are of extreme importance to ensure that images analysis obtains proper results.

In this work we explore an algorithm to detect deteriorated zones due to reflection in dermatological images, either caused by the flash light or uneven illumination conditions. The presented methodology proved to be capable of performing such task. As far as we know this is the first study that performs the pre-processing of dermatological single images captured via mobile phone orientated to the skin cancer prevention.

As future work we intend to expand this study by augmenting the datasets, using images from different mobile phone brands as well as explore the suitability of this algorithm to general macroscopic images. The methodology of classifying the reflection should also be improved, for instance with the inclusion of more subjects in the classification or searching for alternative methods to generate a ground truth dataset.

## Acknowledgements

This work was done under the scope of the project "SMARTSKINS: A Novel Framework for Supervised Mobile Assessment and Risk Triage of Skin Lesion via Non-invasive Screening" with reference PTDC/BBB-BMD/3088/2012 financially supported by Fundação para a Ciência e a Tecnologia in Portugal. The authors would also like to thank the cooperation of the researchers Liliana Ferreira, Filipe Soares, Vânia Guimarães and David Ribeiro in images classification.

# References

[1] H. Kittler, H. Pehamberger and K. Wolff. Binder M.: Diagnostic accuracy of dermoscopy. Lancet Oncol., 3(3):159-165, 2002.

[2] T. Gevers, A. Gijsenij, J. van de Weijer and J-M Geusebroek. Color in Computer Vision: Fundamentals and Applications. Wiley, 2012.

[3] S. Shafer. Using color to separate reflection components. COLOR research and application, 10(4):210–218, 1985.

[4] A. Artusi, F. Banterle and D. Chetverikov. A survey of specularity removal methods. *Computer Graphics Forum*, 30 (8): 2208-2230, Blackwell Publishing Ltd, 2011.

[5] C. Barata, J. S. Marques and J. Rozeira. Detecting the pigment network in dermoscopy images: a directional approach. *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*: 5120-5123, IEEE, 2011

[6] A. DelPozo and S. Savarese. Detecting specular surfaces on natural images. *Computer Vision and Pattern Recognition, 2007, IEEE Conference on:* 1-8, IEEE, 2007.

[7] M. A. Ahmed, F. Pitie and A. Kokaram. Reflection detection in image sequences. *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*: 705-712, IEEE, 2011.

[8] Y. Geng and Z. Jian. Reflections Removal on Mobile Devices. Project Report Stanford University, Digital Image Processing, 2013.

[9] Fraunhofer Portugal AICOS. Melanoma Detection. Internal Project, http://www.fraunhofer.pt/en/fraunhofer_aicos/projects/internal_research/melanoma_detection.html

[10] A. Benoit, A. Caplier, B. Durette and J. Herault. Using Human Visual System Modeling For Bio-Inspired Low Level Image Processing. Computer Vision and Image Understanding, 114: 758-773, Elsevier, 2010.

[11] G. Bradski and A. Kaehler. Learning OpenCV: Computer Vision with the OpenCV Library, O'Reilly Media Inc., 2008.

[12] L. Rosado, M.J. Vasconcelos and M. Ferreira. A Mobile-Based Prototype for Skin Lesion Analysis: Towards a Patient-Oriented Design Approach. *International Journal of Online Engineering* (iJOE), dec 27-29, 2013.

# Unsupervised Restoration of Hair-Occluded Lesion in Dermoscopic Images

Damilola A. Okuboyejo[1]
okuboyejoda@tut.ac.za
Oludayo O. Olugbara[2]
oludayoo@dut.ac.za
Solomon A. Odunaike[1]
odunaikesa@tut.ac.za

[1]Department of Software Engineering. Tshwane University of Technology, South Africa.
[2]Department of Information Technology, Durban University of Technology, South Africa.

**Abstract**

The presence of artefacts such as hair shaft, thin blood vessel, ruler marking and air bubble in medical images makes the diagnosis of skin-related medical images very difficult. This paper uses a two-stage artefact detection termed Fast Image Restoration (FIR) via Canny algorithm and Line Segment Detection (LSD) operation for effective detection of artefacts. The Fast Marching Method (FMM) was applied at each stage for the removal of artefacts from a dermoscopic image in an unsupervised environment while ensuring morphological features of the lesion areas of the image data are preserved. Statistical Analysis performed to determine the accuracy of artefact recognition and repair validation of our method yields a Sensitivity of 98.27%, Specificity of 93.75% and Diagnostic Accuracy of 96.10%. These results indicate that our method gives an acceptable level of accuracy.

# 1 Introduction

Melanoma, which is currently the third prevalent cancer in some parts of the world were for instance reported to occur in 61.7 for every 100000 Australian men and 40.0 for every 100000 Australian women [1]. The development of automated diagnosis systems, capable of performing some level of remote diagnosis of skin cancer diseases such as melanoma and basal cell carcinoma and equally assisting physicians in various imaging task has gained tremendous attention in the bioinformatics and computer vision research [2]. The efforts towards automation of diagnosis procedures are mainly geared to improve the speed of diagnosis and to increase the reproducibility of results. The automated diagnosis has also helped in reducing first-time diagnosis error which sometimes could be as much as 40% [3, 4]. The presence of artefacts such as hair shaft, thin blood vessel, ruler marking and air bubbles in the medical imaging vision makes the diagnosis of skin-related medical imaging very difficult [5-7].

It is important therefore, to mitigate these challenges while still ensuring that quality and time-effective diagnosis can be achieved at a reasonable speed. A popular approach in the literature proposed by Lee [8] is DullRazor hair removal that uses a Top-Hat grayscale morphological closing operation to smoothen out low intensity data of thick dark hairs. The apparent challenge to most of the current effort is the limitation in detecting and repairing

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

thin hair structures while concurrently preserving the morphological features of the image data. The high computational resources required by most of these techniques also presents a challenge.

This study proposes an approach using bilateral filter and a two-stage artefact detection termed Fast Image Restoration (FIR) via Canny algorithm [9] and Line Segment Detection (LSD) [10] operation with fast marching inpainting method in the attempt to autotomize hair occluded artefacts from dermoscopic image in an unsupervised environment while concurrently preserving the morphological features of the lesion image.

# 2 Materials and Methods

## 2.1 Algorithm

The following algorithm is applied in this study for unsupervised restoration of Hair-occluded lesion in a dermoscopic image.

*Step 1*: Input dermoscopic image in RGB colour space.

*Step 2*: Bilateral filtering [11] to perform edge-preserving smoothing operation on the lesion image in a CIE L*a*b space.

*Step 3*: First-level hair-shaft and ruler marking detection using Canny operator [9] with Gaussian filtering and Progressive Probabilistic Hough Transform [12] on result of step 2.

*Step 4*: Apply Fast Marching inpainting Method (FMM) [13].

*Step 5*: Dilate with kernel $k_{2\times2}$ in CIE L*a*b space to remove noise.

*Step 6*: Second-level hair-shaft and ruler marking using a Line segmentation operator.

*Step 7*: Re-applied fast marching inpainting.

*Step 8*: Dilate again with kernel $k_{2\times2}$ in CIE L*a*b space to remove salient salty noise.

## 2.2 Image Standardization

In our experiments, we discovered that CIE*Lab produced a convincing result when compared to its counterparts (CIE L*u*v and CIE X*Y*Z) and the popular YCbCr colour space. This would also help to avoid device dependency requirement of RGB colour space.

## 2.3 Filtering Techniques

### 2.3.1 Gaussian Filter

The Gaussian filter was used to compute the weighted mean of pixels, while being independent of image resolution and spatial location. One major challenge noticed with the use of Gaussian filter is that edge preservation, which is required in detection of the Hair-shaft was an issue [see Figure 1(a), 1(c) and 1(d)].

$$Gaussian\ Filter\ G\langle\delta\rangle_i = \sum_{j\in S} \underbrace{G_\sigma}_{\text{spatial sigma}}(||i-j||)\delta_j \qquad Eq\ 1$$

The problem above is caused by averaging of edges caused by the filters. Bilatering filtering or Diffusion could be used to resolve this challenge.

Figure 1(a) - Original

Figure 1(b) - Bilateral

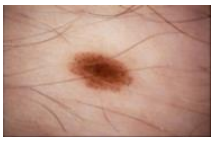


Figure 1(c) - Gaussian

Figure 1(d) - Median

In this study, we applied the Bilateral filtering because of its non-iterative nature and to ensure that smoothening does not stop at thin lines as seen in Anisotropic Diffusion.

### 2.3.2 Bilateral Filter

The Bilateral filtering technique combines both range and domain filters. The traditional low-pass filters perform domain filtering and enforce closeness (neighbourhood in the domain) by weighting pixel values with coefficients that fall off with distance [11]. In contrast, the bilateral filter computes range filtering in addition to domain filtering. When smoothening an image, preservation of edges can be achieved by combining range filtering by domain filtering [see Figure 1(a) and Figure 1(b)]. Our implementation of domain filtering used is according to Gaussian filter [see $Eq$ 1].

$$Gray\ Images: \quad B\langle\delta\rangle_{i_g} = \beta^{-1}\sum_{j_g\in S} G_{\sigma_s}(||i_g - j_g||)G_{\sigma_r}(||i_g - j_g||)\delta_{j_g} \qquad Eq\ 2$$

domain sigma range sigma

$$Colored\ Images: B\langle\delta\rangle_{i_c} = \beta^{-1}\sum_{j_c\in S} G_{\sigma_s}(||i_c - j_c||)G_{\sigma_r}(||i_c - j_c||)\delta_{j_c} \qquad Eq\ 3$$

domain sigma range sigma

$\beta^{-1}$ *is the normalization factor*

$\left|\left|i_g - j_g\right|\right|$ *is the intensity changes*

$||i_c - j_c||$*is the color difference in either RGB or* CIE $La^*b^*$

$\delta_{j_g}$*and* $\delta_{j_c}$*represents scalar and 3D vector quantity respectively*

## 2.4 Dilation Morphological Operation

In relation to mathematical morphology, dilation provide a maximizing operation technique for causing bright regions within an image to grow, thus leading to removal of salient noise. It works by convolving an image $\delta$ with a structuring element (kernel) $\kappa$ of any shape or size.

$$\delta \oplus \kappa = \bigcup(\kappa + \iota; \iota \in \delta) \qquad Eq\ 4$$

## 2.5 Artefact Recognition

### 2.5.1 Progressive Probabilistic Hough Transform

The Progressive Probabilistic Hough Transform (PPHT) was applied to estimate hair-shafts and ruler marking based on the result of the Canny edge [see Figure 2(c) - 2(e)]. The PPHT performs line segmentation by using the difference in the fraction of votes needed to effectively detect lines with different numbers of supporting points resulting from operators such as Sobel and Canny.

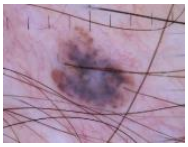


Figure 2(a) - source image

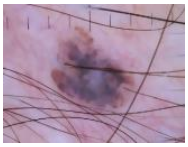


Figure 2(b) – filtered image

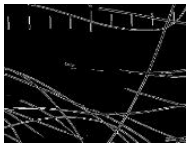


Figure 2(c) – canny edge

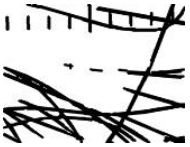


Figure 2(d) – mask from PPHT

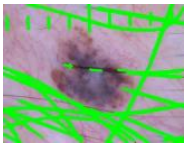


Figure 2(e) – line segment by PPHT

## 2.5.2 Fast Marching Inpainting

Inpainting generally involves using textures and colours identical to a defined boundary of an identified gap to fill different regions in the gap defined by contour lines. We applied the Fast Marching Methods (FMM) proposed by Talea [13] as highlighted in the pseudo code below.

*Let* $\delta$ *represent an image having gradient* $\mathfrak{g}$, such that
$\delta_i = pixel\ value\ of\ image\ \delta\ at\ i_{x,y}$
$\delta_j = pixel\ value\ of\ image\ \delta\ at\ j_{x,y}$
$\mathfrak{g}_i = gradient\ value\ of\ \delta_i$ ; $\mathfrak{g}_j = gradient\ value\ of\ \delta_j$

*if* $\mathfrak{D}$ *represent First Order Approximation* $(FDA)$, such that
$\mathfrak{D}_i = FDA\ of\ \delta_i$
$\mathfrak{D}_j = FDA\ of\ \delta_j$
then, it follows that
$\mathfrak{D}_i = \delta_j + \mathfrak{g}_j(i-j)$
*let* $\mathcal{N}_{z_i} = neighborhood\ pixel\ of\ size\ z\ of\ the\ known\ image\ around\ i$
$\beta_{i,j} = weighted\ function\ to\ propagate\ sharp\ image\ details\ to\ inapainted\ arena$

*if* $\Re$ *represent region to be inpainted*
$\mathbb{B}$ *represent boundary of the region*
$\Re_i$ *represent initial position in the region*
$\xi_i$ *represent inpaint point* $i$ ; $\xi_i = \frac{\sum_{j \in \mathcal{N}_{z_i}} \beta_{i,j} \mathfrak{D}_i}{\sum_{j \in \mathcal{N}_{z_i}} \beta_{i,j}}$ $Eq\ 5$

While ($\mathbb{B}$ still contains valid points)
$i$ = pixel $\wp$ nearest to $\Re_i$
perform interpolation on $i$ using $Eq$ 5
advance $\mathbb{B}$ into $\Re$

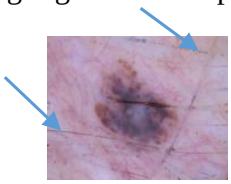


Figure 3 – Inpainted image after applying FMM

While the combination of Bilateral filtering, Canny operation and PPHT using well-defined hysteresis threshold achieved a level of success, a thorough detection of the noise was not performed [see sample arrows drawn in Figure 3]. This challenge was addressed by performing a second-level artefact recognition using the line segmentation operator.

## 2.5.3 Line Segment Detector

We applied the LSD technique on results produced by FMM [see Figure 3] and generated edge points for the remaining hair-shafts as highlighted in Figure 4(a). The mask of Figure 4(a) was obtained and used by FMM to perform a second-level inpainting that resulted in Figure 4(c). In this paper, we followed the recommendations by Gioi et al.[10] considering our objectives. The approach proposed in [10] is based on [14, 15] and Desolneux's line validation method [16] using Helmholtz principle [17].

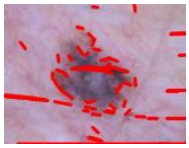


Figure 4(a) – line segment by LSD

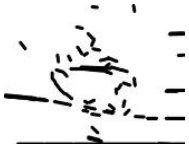


Figure 4(b) – LSD mask

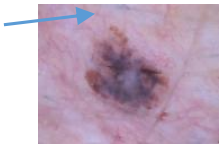


Figure 4(c) – result after applying FMM

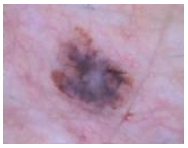


Figure 4(d) – morphologically dilated image

While the result in Figure 4(c) looks promising, some artefacts as pointed out by the arrows were still noticeable. Morphological dilation was performed on Figure 4(c) to remove the salient salty noise and this resulted in output in Figure 4(d).

# 3 Results and Discussions

Our method has been applied to a dataset [1] of a total of 299 images. A total of 294 of the images are melanocytic lesion bound and 5 are non-melanocytic. Out of the 294 melanocytic images, 105 images were already histopathologically certified by the Dermatologists as melanoma and the remaining 189 images as being benign naevus. The ground truth of the artefacts was obtained from manual identification and sketch-trace of the visible artefacts in a given image.

The implementation of each step in the highlighted algorithm of section 2.1 was computed on a virtual machine running Ubuntu 12.0.4.3 LTS, with base memory of 3096MB and 2 processors. The machine has VT-X/AMD-V hardware virtualization capability with nested paging. Average speed of execution of the complete algorithm is 380ms.

In order to ensure that hair-shaft, including fine hairs and other related artefacts are well recognized, the first part of our approach is centred on smoothening of each image while retaining the edges of artefacts to be repaired. Our study indicates that the best filtering technique for our objective is bilateral filter [see Figure 1(a) – 1(d)]. The second part of our approach is focused on identification of artefacts in a given lesion using a two-stage detection process of the Canny operator with PPHT and line segment detector in that order. A remarkable result was achieved in our study because of the use of both Canny and LSD, rather than using only either of them. It is important therefore to state that applying the Canny operator with PPHT before LSD produced a better result than the reverse.

The statistical analysis based on the metrics of Sensitivity, Specificity and Diagnostic Accuracy was used to determine the performance of artefact recognition and repair validation. Our proposed approach reports a true positive rate (Sensitivity) of 98.27% and a true negative rate (Specificity) of 93.75%. The diagnostic accuracy achieved is recorded at level high of 96.10%. Figure 5(a) – Figure 5(d) and Figure 6(a) – Figure 6(d) highlight a comparable output of our approach with that of DullRazor [8] and Zhou et al. [7].

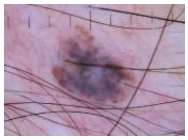


Figure 5(a) – Original Image

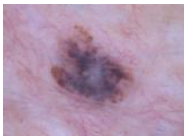


Figure 5(b) – DullRazor

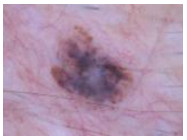


Figure 5(c) – Zhou

Figure 5(d) – Our approach

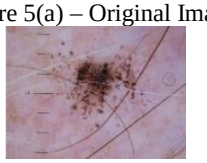


Figure 6(a) – Original Image

Figure 6(b) – DullRazor

Figure 6(c) – Zhou

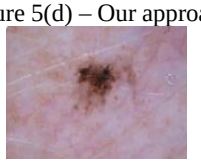


Figure 6(d) – Our approach

# 4 Conclusion

We proposed in this study a fast and effective approach towards repairing hair-occluded and other artefacts in a dermoscopic images. The artefact recognition involved a two-stage process of using Canny and LSD operator. Experimental results indicate our proposed approach is dependable, robust and can effectively repair lesions occluded with hair-shafts and air bubbles. Dermatologists can easily use our approach as a form of pre-processing of dermoscopic images before lesion segmentation and classification. Our approach worked well on various types of lesion as described in section 3.

[1] The images were provided as a support in part by Dermatology Society of South Africa (DSSA).

# References

[**1**] Australian Institute of Health and Welfare & Australian Association of Cancer Registeries 2012, Cancer in Australia, in, 2012.

[**2**] R. Dobrescu, M. Dobrescu, S. Mocanu, D. Popescu, Medical images classification for skin cancer diagnosis based on combined texture and fractal analysis.Proceedings of WSEAS Transactions on Biology and Biomedicine; 2010. July 2010.

[**3**] M. Kantrowitz. Pathology Reports May Contain Errors. 2009. Available from: http://www.kantrowitz.com/cancerpoints/diagnosiserrors.html (Accessed Jan 16th, 2014).

[**4**] H. Singh, S. Sethi, M. Raber, L.A. Petersen, Errors in Cancer Diagnosis: Current Understanding and Future Directions, JCO, (2007) 5009-5018.

[**5**] Q. Abbas, I.F. García, M. Rashid, Unsupervised skin lesions border detection via two-dimensional image analysis, Comput Methods Program Biomed, (2010).

[**6**] A.K. Mittra, R. Parekh, Automated detection of skin disease using texture features, International Journal of Engineering Science and Technology, 3 (2011).

[**7**] H. Zhou, M. Chen, R. Gass, J.M. Rehg, L. Ferris, J. Ho, L. Drogowski, Feature Preserving Artifact Removal from Dermoscopy Images.Proceedings of SPIE Medical Imaging: Image Processing; 2008. San Diego, USA.

[**8**] K.T. Lee, R. Galagher, A. Coldman, D. Mclean, Dullrazor: A software approach to hair removal from images, Comput Biol Med, 27 (1997) 533-543.

[**9**] J. Canny, A Computational Approach to Edge Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol PAM 1-8, No 6. (1986).

[**10**] R.G.v. Gioi, J. Jakubowicz, J.-M. Morcel, G. Randall, LSD: A Line Segment Detectot, Image Processing On Line, (2012).

[**11**] C. Tomasi, R. Manduchi, Bilateral Filtering for Gray and Color Images.Proceedings of IEEE International Conference on Computer Vision; 1998. Bombay, India.

[**12**] J. Matas, C. Galambos, J. Kittler, Robust Detection of Lines Using the Progressive Probabilistic Hough Transform, Computer Vision and Image Understanding, 78 (2000) 119-137.

[**13**] A.C. Talea, An Image Inpainting Technique based on the Fast Marching Method, Journal of Graphics Tools, Vol 9. No. 1 (2004) 25-36.

[**14**] J.B. Burns, A.R. Hanson, E.M. Rieseman, Extracting Straight Lines, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 4 (1986) 425-455.

[**15**] J.R. Beveridge, C. Graves, C. Lesher, Some Lessons learned from Coding the Burns Line Extraction Algorithm in the Darpa Image Understanding Environment, Technical Report CS-96-125C.S. Department, Colorado State University, 1996. Available from: http://www.cs.colostate.edu/TechReports/Reports/1996/tr96-125.pdf (Accessed: February 2, 2014).

[**16**] A. Desolneux, L. Moisan, J.-M. Morel, Meaningful Alignments, International Journal of Computer Vision, vol. 40, no. 1 (2000) 7-23.

[**17**] A. Desolneux, L. Moisan, J.-M. Morel, From Gestalt Theory to Image Analysis - A Probabilistic Approach, International Applied Mathematics, vol. 34 (2008).

# Segmentation

# A Hybrid Method for Haemorrhage Segmentation in Trauma Brain CT

Mohammadreza Soltaninejad[1]
msoltaninejad@lincoln.ac.uk
Tryphon Lambrou[1]
tlambrou@lincoln.ac.uk
Adnan Qureshi[2]
adnan.qureshi@beds.ac.uk
Nigel Allinson[1]
nallinson11@gmail.com
Xujiong Ye[1]
xye@lincoln.ac.uk

[1] School of Computer Science, University of Lincoln, UK
[2] Department of Computer Science and Technology, University of Bedfordshire, UK

**Abstract**

Traumatic brain injuries are important causes of disability and death. Physicians use CT or MRI images to observe the trauma and measure its severity for diagnosis and treatment. Due to the overlap of haemorrhage and normal brain tissues, segmentation methods sometimes lead to false results. In this paper, we present a hybrid method to segment the haemorrhage region in trauma brain CT images. Firstly, the images are partitioned to small segments called superpixels and supervoxels in 2D and 3D spaces, respectively. Then the haemorrhage superpixels/supervoxels are grouped using their average intensity as feature. Finally, a distance regularized level-set is used to accurately delineate the exact boundary of the haemorrhage region. Evaluation is performed using the Jaccard overlap measure of our proposed technique against a modified distance regularized level-set and against the manually segmented ground truth. Our results suggest that performing level-set after superpixel/supervoxel segmentation provides better segmentation than superpixel/supervoxel intensity grouping alone and both these schemes perform better than the modified distance regularized level-set evolution method.

**Keywords**, Traumatic brain injury, brain CT images, superpixel/supervoxel segmentation, level-set.

# 1 Introduction

Traumatic brain injuries (TBI) are usually caused by external forces on brain in accidents or head strikes and are considered an important cause of disabilities and death. TBI can range from subdural haematoma (SDH), epidural haematoma (EDH), intra-cerebral haemorrhage (ICH), cerebral contusion, diffuse axonal injury (DAI) and cerebellar haemorrhage. Medical imaging modalities such as CT are widely used to establish severity and diagnosis and decide targeted therapy within early hours of injuries

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

and an accurate and timely decision is imperative for physicians to manage the patient more precisely and ensure better prognosis and outcomes. However, manually assessing neuro-imaging studies is a time consuming process and in such cases, computerized processing techniques are a useful 'second opinion' for physicians to measure the size, location and extent of traumatic region(s) more succinctly.

Several methods have been used to segment and quantified the haemorrhage regions in brain CT images. Bhaudauria et al [1] proposed combination of Fuzzy C-Means clustering and region-based active contour to segment the haemorrhage region. They used clustering to find an initial mask and then propagate it to the haemorrhage boundaries using region-based active contour. Prakash et al. [2] proposed an automatic procedure based on a modified distance regularized level set evolution (DRLSE) to segment haemorrhage in brain CT images. The method used DRLSE twice with different parameters. Firstly, a general intensity mask is defined by a set of DRLSE parameters to localize the haemorrhage region rapidly. Then the boundaries are detected in another level-set step by choosing the parameters so that the more accurate boundaries are detected. This method is sensitive to initial threshold values. Tao [3] developed a method to identify small acute intracranial haemorrhage. The probable regions are detected based on left-right asymmetry regarding to the vertical asymmetry line.

Since the intensity values for haemorrhages overlaps with the normal tissues, it causes difficulties for segmentation task. The method proposed in [2] results in leakages in boundaries in which the normal brain tissue and trauma have very close intensity values. Furthermore, global threshold values are different for individual cases, and the user should predetermine them. In this paper, we intend to segment the haemorrhage regions using a hybrid method consisting of superpixel (SP) /supervoxel (SV) segmentation partitioning, grouping and level-set boundary detection. The method has been evaluated in both 2D and 3D spaces. 3D volume of segmented haemorrhage is calculated and then compared to the ground truth.

The paper is organized as follows. Section 2 describes the steps of the proposed method. Section 3 analyses experimental results and simulations. Finally, section 4 presents conclusion and future work.

# 2 Proposed Method

The proposed method is a hybrid algorithm and consists of three main steps and is implemented in two dimensional and three dimensional spaces. The flowchart of the steps for the proposed method is shown in Figure 1.

Figure 1: Flowchart of the proposed method in 2D, the 3D implementation is similar.

## 2.1 Brain segmentation

Scalp and background are easily segmented by thresholding due to their distinct Hounsfield Unit values. The soft tissue which contains both brain and traumatic regions are segmented using conventional grey-level thresholding. The outcome consists of all types of soft tissues such as brain, skin etc. The connected components of the segmented regions are extracted and the largest 3D component is considered as brain region. The separated brain image is shown in Figure 2-a.

## 2.2 Superpixel Segmentation

We use the simple linear iterative clustering (SLIC) superpixel method proposed in [4] for segmentation. Each slice is gridded to small equal square sections of a user-defined size, which are considered as initial superpixels. To perform the process we need the centres of each region for further calculations. In the first instance, the geometrical centres of these squares are considered as superpixel region centres. Then the locations of the centres are updated according to the mean value of the pixel coordinates of each cluster. The distance between each pixel in the image to the bounded cluster centres are computed and a label of the closest cluster centre is assigned to the target pixel. The distance is calculated by defining the intensity difference between the *i*th pixel and the *j*th pixel, $d_c$, and location distance, $d_s$, with the following formulas:

$$d_c = \sqrt{(I_j - I_i)^2} \tag{1}$$

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2 + (z_j - z_i)^2} \tag{2}$$

where $I$ is intensity and $(x, y, z)$ are pixel location coordinates in the image. By combining the distances and applying a compactness factor, $m$, we derive the following distance measure:

$$D = \sqrt{d_c^2 + \left(\frac{d_s}{s}\right)^2 m^2} \tag{3}$$

where, $m$, is the compactness coefficient. A higher value of $m$ results in more compact segments and a lower value creates more flexible boundaries. The principles of three-dimensional supervoxel are similar to 2D segmentation. There are only a few modifications in initial gridding and the formulation. For the specific case of brain CT images, the intensity distance is the same, because the images are grey-level. Only the z direction is added to the location distance, $d_s$. Figure 2.b shows the superpixel segmentation of a CT brain image containing a haemorrhage region.

## 2.3 Grouping the Superpixels

A superpixel feature map is created by calculating the mean intensity of all the pixels within each superpixel. Figure 2.c shows an example of the superpixel feature map. Once the user selects a point inside the target region, the corresponding intensity feature related to its corresponding superpixel is determined. All the superpixels with a value in the range of a confidence margin 3, adjacent to the chosen superpixel are selected. The neighbouring superpixels are determined using connected components by finding the adjacent labels connected to the central superpixel. Finally, these superpixels are grouped together to make an initial contour for the next stage of the proposed method. The outcome of

superpixel grouping is shown in Figure 2.d. A similar three dimensional approach was used in the case of supervoxel implementation.

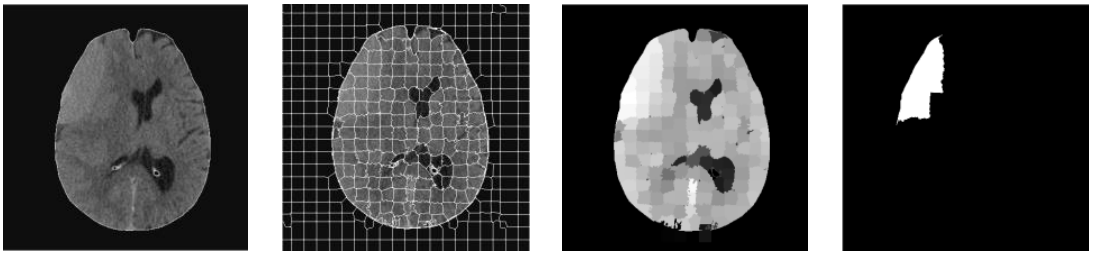
## 2.4 Level-set Segmentation

The level-set (LS) algorithm tries to move the contour of a zero level set function to the image boundaries by minimizing a cost function. Li et al [5] proposed a new approach for the level set method which increases the stability and omit the reinitialization and its numerical errors. Their method tries to move the initial level set to the boundaries by minimizing an energy function:

$$\varepsilon(\varphi) = \mu R_P(\varphi) + \varepsilon_{ext}(\varphi)$$

where, $R_P(\varphi)$ is the level-set regularization term, $\mu$ is a constant, and $\varepsilon_{ext}(\varphi)$ is the external energy of the image.

The region obtained in the superpixel grouping stage is used as an initial level set function for running DRLSE. For noise reduction, a 3x3 median filter is applied on the image. This guarantees smoothing without affecting the edges in the image and missing any boundary information. The level-set parameters are empirically defined as follows: weight of distance regularization term = 0.4, the coefficient of the weighted length term L = 0.5, coefficient of the weighted area term A = 0.1, the width of the Dirac delta function = 1.5, internal iterations = 5, maximum external iteration = 100. The stopping condition for the iterations is based on the difference of pixel number between two iterations. Regarding to [4] this threshold is empirically defined as 10 pixels. This step guarantees that the boundaries of the superpixel segmentation stick to the haemorrhage region boundaries.



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

Figure 2: Superpixel segmentation and grouping result: a) brain extracted from original image, b) superpixel segmentation result, c) feature map, d) grouping the partitions.

# 3 Results

The implementation is utilised in MATLAB 2013b installed on a PC with Windows 8 platform. The proposed method has been evaluated on a cohort of seven CT scans containing different types of traumatic injuries such as: intra cerebellar, Gyrriform, cerebral contusion, and intracranial haemorrhage. The data were acquired on a Toshiba, Activion 16, CT scan machine. Each CT slice is stored in DICOM file format, with matrix size 512x512, and spatial resolution 0.468x0.468 mm, and slice thickness of 1.0 mm. Overall, 398 haemorrhage affected CT slices were examined; analytically the number of affected slices for subjects 1 to 7 were: 90, 87, 84, 26, 20, 31, 60, respectively. The ground truths for evaluation of the results are manual segmentations by an experienced

neurologist. The Jaccard measure is used to evaluate the overlap between the segmentation results and the manual ground truth:

$$J = \frac{|M \cap S|}{|M \cup S|} \qquad (4)$$

where, $M$ and $S$ are the manual and the proposed segmentation masks, respectively.

The segmentation results for both 2D and 3D procedures are compared with the results produced by the modified DRLSE method, superpixel and supervoxel grouping from step 2.3.

The compactness coefficient for both 2D and 3D SLIC segmentation is considered 10. The resolution for initial superpixels is 10x10 and for the supervoxel volumes is 10x10x5. Figure 3a – 3c, shows the boundaries extracted using our hybrid supervoxel combined with level-set segmentation (blue line) and manual segmentation (red line). Three-dimensional rendering of the proposed method, namely superpixel segmentation followed by level-set, is presented in Figure 3d.

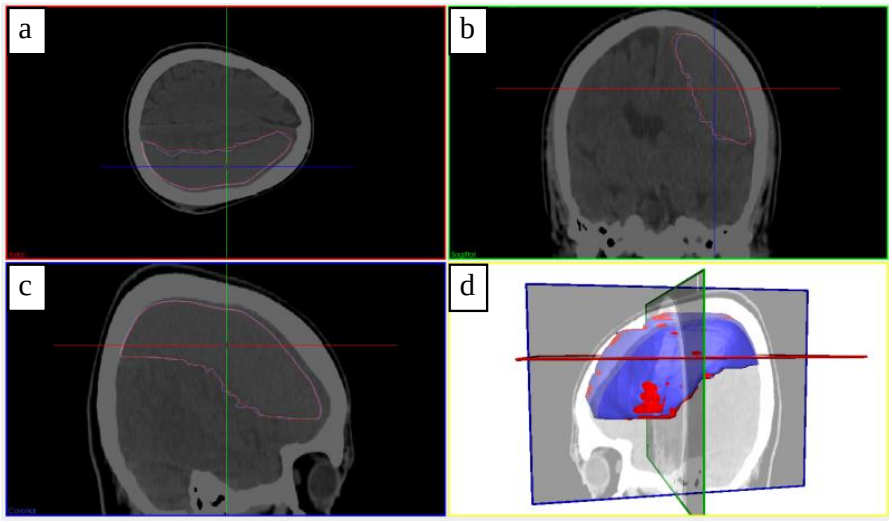



Figure 3a – 3c, axial, coronal, and sagittal views respectively. Figure 3d, shows a 3D rendering with proposed method and manual overlays.

The mean and standard deviation of Jaccard overlap measure are listed in Table 1 for each dataset. The results show that the proposed method improves the accuracy of the segmentation comparing to modified DRLSE method. Furthermore, the secondary stage of level-set segmentation for both superpixel and supervoxel partitions, increase the accuracy in all the datasets.

# 4 Conclusion

A hybrid haemorrhage segmentation and measurement technique is proposed in this paper. The method is based on superpixel/supervoxel grouping and subsequent level-set segmentation. Subvolume partitions with similar features to the selected region point, that are connected together, are grouped as initial volumes and then an accurate boundary is

determined using level-set method. The results show that the proposed technique offers better robustness in older trauma cases were CT HU of the affected intensity values are near to those of normal tissues. The segmentation accuracy in the case of small haemorrhage regions is not satisfactory; however, we solved this problem by decreasing the initial superpixel/supervoxel size. In future works, other measures such as textural features of the superpixels/supervoxels will be considered for their groupings.

| | **Level-set** Mean ± STD | **Superpixel Segmentation** Mean ± STD | **Supervoxel Segmentation** Mean ± STD | **Superpixel + Level-set** Mean ± STD | **Supervoxel + Level-set** Mean ± STD |
|---|---|---|---|---|---|
| **Case 1** | 0.6716 ± 0.1201 | 0.8328 ± 0.0797 | 0.7862 ± 0.1136 | 0.8621 ± 0.0701 | 0.8161 ± 0.1017 |
| **Case 2** | 0.3242 ± 0.2218 | 0.7763 ± 0.0915 | 0.7367 ± 0.1454 | 0.8102 ± 0.0888 | 0.7586 ± 0.1407 |
| **Case 3** | 0.4605 ± 0.2577 | 0.7962 ± 0.1368 | 0.7311 ± 0.2250 | 0.8158 ± 0.1109 | 0.7498 ± 0.2296 |
| **Case 4** | 0.6397 ± 0.3364 | 0.7987 ± 0.0562 | 0.6373 ± 0.1459 | 0.8469 ± 0.0450 | 0.6806 ± 0.1632 |
| **Case 5** | 0.5081 ± 0.1873 | 0.8572 ± 0.0829 | 0.8193 ± 0.0760 | 0.8723 ± 0.0775 | 0.8383 ± 0.0784 |
| **Case 6** | 0.4565 ± 0.1009 | 0.6511 ± 0.0750 | 0.6237 ± 0.1045 | 0.6604 ± 0.0878 | 0.6388 ± 0.1213 |
| **Case 7** | 0.5642 ± 0.1862 | 0.7985 ± 0.0760 | 0.7573 ± 0.0659 | 0.7810 ± 0.1002 | 0.7827 ± 0.0630 |

Table 1: Mean and standard deviation of the Jaccard overlap measure.

# Acknowledgment

This research was supported by European FP7 collaborative project “MyHealthAvatar” (GA No: 600929).

# References

[1] H.S. Bhadauria, A. Singh, and M.L. Dewal. An integrated method for hemorrhage segmentation from brain CT imaging. *Computers and Electrical Engineering*, 39: 1527–1536, 2013.

[2] K. N. Bhanu Prakash, S. Zhou, T. C. Morgan, D. F. Hanley, and W. L. Nowinski. Segmentation and quantification of intra-ventricular/cerebral hemorrhage in CT scans by modified distance regularized level set evolution technique, *International Journal of Computer Assisted Radiology and Surgery*, 7(5): 785–798, 2012.

[3] T. Chan. Computer aided detection of small acute intracranial hemorrhage on computer tomography of brain, *Computerized Medical Imaging and Graphics*, 31: 285–298, 2007.

[4] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34 (11): 2274-2282, 2012.

[5] C. Li, C. Xu, C. Gui, M. D. Fox. Distance Regularized Level Set Evolution and Its Application to Image Segmentation, *IEEE Transactions on Image Processing*, 19 (12): 3243-3254, 2010.

# Ensemble registration: Combining groupwise registration and segmentation

Sri Purwani[1,2]
sri.purwani@postgrad.manchester.ac.uk
Tim Cootes[1]
t.cootes@manchester.ac.uk
Carole Twining[1]
carole.twining@manchester.ac.uk

[1] Imaging Science,
The University of Manchester, UK
[2] Mathematics Department,
Faculty of Mathematics and
Natural Sciences,
Padjadjaran University, Indonesia

**Abstract**

We investigate incorporating structural information from segmentation into a groupwise registration framework. Previous work by Petrovic et al., using MR brain images, showed that using tissue fractions to help construct an intensity reference image gives better results than just using intensity images alone. In their work, a Gaussian Mixture Model (GMM) was fitted to the 1D intensity histogram, then used to construct tissue fraction images for each example. The mean fraction images were then used to create an artificial intensity reference for the registration.

By using only the mean, this discarded much of the structural information. We retain all this information, and augment each intensity image with its set of tissue fraction images (and also intensity gradient images) to form an image ensemble for each example. We then perform groupwise registration using these ensembles of images.

This groupwise ensemble registration is applied to the same real-world dataset as used by Petrovic et al. Ground-truth labels enable quantitative evaluation to be performed. It is shown that ensemble registration gives quantitatively better results than the algorithm of Petrovic et al., and that the best results are achieved when more than one of the three types of images (intensity, tissue fraction and gradient) are included as an ensemble.

# 1 Introduction

Previous work [1,6,7] has shown that integrating segmentation and registration in a unified framework gives better results. Asburner et al. [1] showed that repeated iteration among segmentation, bias correction, and registration of brain images gives better results than serial applications of each component. On the other hand, Konukoglu et al. [6] integrated segmentation into registration to improve robustness of the registration. The segmentation automatically labels structures of human anatomies in its found regions. This integration shows improvement on CT image alignments.

In Petrovic et al. [7], a Gaussian Mixture Model (GMM) is fitted to the intensity distribution of MR brain images. Each image is taken to consist of three main tissues (CSF, grey matter (GM) and white matter (WM)) plus partial-volume voxels. This model then enables the construction of a set of tissue fraction images for each image in the set. The mean fraction images across the group are then used to create an artificial intensity

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

reference to be registered with its original intensity image during registration. Their results out-performed those which just used intensity images alone.

However, the use of a 1D intensity image histogram in their method sometimes suffers from the problem of failing to find the correct parameters, such as the tissue mean intensities. Hence segmentation fails for some images, as is shown in Figure 1(a). However, since just the ***mean*** fraction images were used to create the reference, their algorithm was not overly sensitive to these failures. However, this also means that only a limited amount of the structural information is actually used.

Our current method is designed to use all of the structural information available. This means including not just the tissue fraction images, but also intensity and gradient images to form an ensemble of images for registration. This integration hence makes better use of all available information for alignment of structures.

However, this requires a more robust GMM fitting algorithm than that used by Petrovic et al. [7]. We achieve this by using the 2D intensity value-and-gradient histogram [2], which gives a more robust estimate of the GMM intensity parameters. We show that our method improves on that of Petrovic et al. [7]. Tanimoto overlap values [3] for the ground truth labels are used to evaluate the registration, and show that combining image types into an ensemble gives the best results.

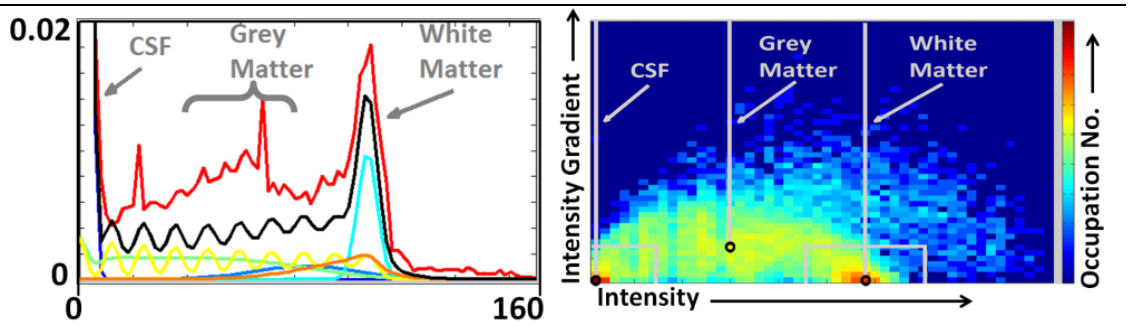
# 2 Method

Given a set of images, we are going to perform groupwise ensemble registration from each example image in the set. The way to do this is by segmenting each image into tissue fraction images and then constructing an ensemble of images consisting of those fraction images and the original intensity image (plus intensity gradient image derived from the original image). This can be viewed in two ways: either as each example consisting of a stack (or ensemble) of images, or as each example image now being a vector-valued feature image, with intensity, fraction, gradient etc. being components of the vector. This will be further explained in the following sections.

## 2.1 Tissue Segmentation

We consider the case where different tissues in an image have different intensity distributions hence tissue segmentation can be performed by analysis of the intensity histogram. In Figure 1(a), we show a typical 1D intensity histogram (red). The GMM fitting algorithm used by Petrovic et al. [7] gives the model shown in black. In this case the 1D algorithm ***fails*** to correctly locate the pure-grey-matter Gaussian, and the segmentation fails. We hence use the 2D intensity value-and-gradient histogram instead [2]. As can be seen in Figure 1(b), the pure tissue peaks are located at the feet of the arcs that correspond to partial-volume mixtures of tissues. The gradient, which describes intensity changes within images, is highest at the boundary between tissues, hence can separate partial-volume boundary pixels from the pure-tissue pixels (see [2] for further details). So it could be said that as well as gradient information being useful for registration [5], this shows that it is useful for segmentation as well.

A greedy algorithm is used to locate the peaks/means for model fitting (Figure 1(b)). After finding the first peak, this algorithm defines an excluded zone (grey lines) before finding the next peak. This is to avoid having multiple local peaks, and is repeated until all




(a) 1D histogram (b) 2D histogram

Figure 1: Intensity 1D histogram (Left) and the new intensity-gradient 2D histogram (Right), to the same horizontal scale. (a) The red line is the image histogram, and the black line is the ***failed*** GMM fit, which consists of a sum over 3 pure tissue Gaussians, plus 3 mixtures (coloured lines); (b) The 2D Intensity-gradient vs Intensity histogram. The yellow/red regions at the feet of the yellow arcs give the new estimates of the pure-tissue means (shown by the vertical grey lines and the black circles).

three peaks are found. These peaks give the means for the three Gaussians, which are kept fixed during the optimisation of the fit, and we optimise all the other parameters such as variance and weights by using Sum of Squared Differences.

We have found that using this procedure, it was possible to correctly fit all the image histograms in the set, including those where the 1D histogram fitting procedure is failed.

## 2.2 Ensemble Registration

Now we have robust segmentation of the histogram, this can be used to relate image intensities to tissue content [2], hence segment the image to obtain 3 tissue-fraction images. These can then be combined with the original intensity image. As noted above, we also include the intensity gradient image [5]. This then gives a 5 component image ensemble for each example in the set. Registration can then be performed to the reference in the usual way, but with a separate reference image for each component. The objective function then must be summed across components as well as across the set. Using Sum of Absolute Differences, we then have an objective function of the form:

$$\mathcal{L} = \sum_{\alpha=1}^{n} \sum_{j=1}^{m} \sum_{A} |I_{\alpha}^{j}(x_A) - \tilde{I}_{ref}^{j}(x_A)| \qquad \textbf{(1)}$$

where $\alpha = 1, 2, \ldots, n$ , $j = 1, 2, \ldots, m$ are the indices for example number from the set, and image ensemble component respectively, whereas A runs over all pixels in the image. For each component, $\tilde{I}_{ref}^{j}$ is the reference image for that component, and is taken to be the mean over the set. Each image in the ensemble is scaled to lie in the range 0 to 1, with the full range of values retained, so that the different components are commensurate. Each component hence contributes equally to the objective function. It would obviously also be possible to weight the individual components, but here we are just interested in using a component compared to not using it, hence we give all components equal weight.

The same multi-resolution image warping framework (piece-wise affine deformation) as was used in [7] is applied, to enable a true comparison of the registration objective functions. The algorithm is applied to the same dataset as in [7] (37 MR brain image slices, from 37 different individuals, of size $190 \times 190$ pixels).

### 2.3 Evaluation: Ground Truth Labels

On this set of MR brain images there are 64 ground truth tissue labels. However, some labels are of very small volume. Hence to get a more informative comparison, Petrovic et al. [7] consider just the main labels having the most contribution to the image set, which are the 8 labels consisting of white matter, grey matter, caudate, and lateral ventricle, left and right in each case. For analysis and comparison purposes with other datasets, we arrange our labels into 4 tissue sets, CSF (not including background), GM, WM, and other tissues. Some samples of ground truth labels from this image set are given in Figure 2.

We perform quantitative evaluation of the registration using Tanimoto overlaps on these different tissue sets, which we found to be more informative than the single variously-weighted combination used in [7]. These evaluation results are presented in the experiment section.

Figure 2: The main ground truth labels.

## 3 Experiments

For our initial comparisons, we consider two modifications to the original method of Petrovic et al. [7] (results in Table 1). It was noted that they had three pairwise mixtures in their method, including the mixture WM-CSF. Given the problems with segmentation (see Section 2.1), we first restricted mixtures to those sorted by intensity[1], so we include CSF-GM, and GM-WM, but ***not*** CSF-WM [4]. Together with pure tissues, this then gives 5 possible classes. We also modify slightly the mappings from intensity values to fraction values. In [7] (see Eqn. (4)), a hard assignment of pixels to the most probable class was made, whereas we use a softer assignment [4], which gives a smoother, continuous mapping. In the second modification, we additionally replace the 1D histogram method with the 2D histogram method (see Section 2.1). As well as these slight modifications of the original algorithm, we also include the Ensemble registration (1). For completeness, we consider all seven possible combinations of components in the ensemble, where F denotes the inclusion of the 3 fraction images, I denotes intensity images, and G intensity gradient.

It is slightly surprising that the first two modifications, which could be viewed as having a better theoretical justification, either make things worse, or have a negligible effect. More importantly, it can be seen that the Ensemble registration can produce slightly better results than the algorithm of Petrovic et al. [7]. In terms of single-component-type registration, in general gradient alone is the worst, followed by intensity, but with fraction images being the best. This is perhaps explained by the fact that the single-tissue mean intensities vary for this dataset (which is why the example-specific reconstructed intensity reference images were created by Petrovic et al. originally). This might mean that gradient (which removes this factor) should perform better than intensity is. Except the largest gradients are in the region of the skull (see Figure 3), but there are no labels (see Figure 2) associated with the skull, so that this region does not even enter into the evaluation!

[1] See Figure 1b), where the two prominent yellow arcs correspond to the two included mixtures, with little evidence of a CSF-WM arc (the other arc corresponds to skull mixtures, which are not included in the GMM).

The pairs-of-component-types results respect this ordering, so that fraction-plus-gradient outperforms intensity-plus-gradient, intensity-plus-fraction outperforms intensity plus-gradient, and fraction-with-intensity is better than fraction-with-gradient. Finally, the combination of all three component types is always ranked first or second.

These results clearly show that the explicit structural information in the tissue fraction images is needed, and that it seems to be of better value than simpler quantities such as intensity gradient (since we take the absolute magnitude of the gradient, rather than a full vector-valued gradient image, some information is lost when we take the gradient).

Figure 3 shows initial and final component reference images respectively for the full ensemble registration (FIG). This shows that as registration progresses the alignment of structures improve, and the reference images consequently become sharper for all 5 components of the ensemble.

| Registration | Tanimoto Values for Tissue Set | | | |
|---|---|---|---|---|
| | CSF | grey | white | others |
| Petrovic et al. [7] | 0.6854(9) | 0.5804(6) | 0.700(1) | 0.458(3) |
| Modification 1 | 0.665(4) | 0.576(1) | 0.6884(7) | 0.457(7) |
| Modification 2 | 0.675(2) | 0.571(2) | 0.698(1) | 0.459(4) |
| F | 0.678(4) | 0.575(1) | 0.7021(8) | 0.507(4) |
| I | 0.646(3) | 0.5393(7) | 0.652(3) | 0.32(1) |
| G | 0.524(4) | 0.468(1) | 0.6211(9) | 0.499(9) |
| I F | **0.690(5)** | **0.582(1)** | **0.7069(1)** | 0.508(5) |
| G F | 0.675(5) | 0.577(1) | 0.704(1) | **0.515(4)** |
| G I | 0.666(2) | 0.553(1) | 0.6792(6) | 0.462(4) |
| F I G | **0.686(3)** | **0.584(1)** | **0.709(1)** | **0.509(2)** |

Table 1: Tanimoto values by tissue set, for various registration methods. The best two values in each column are highlighted. Bracketed numerals show the estimated errors.

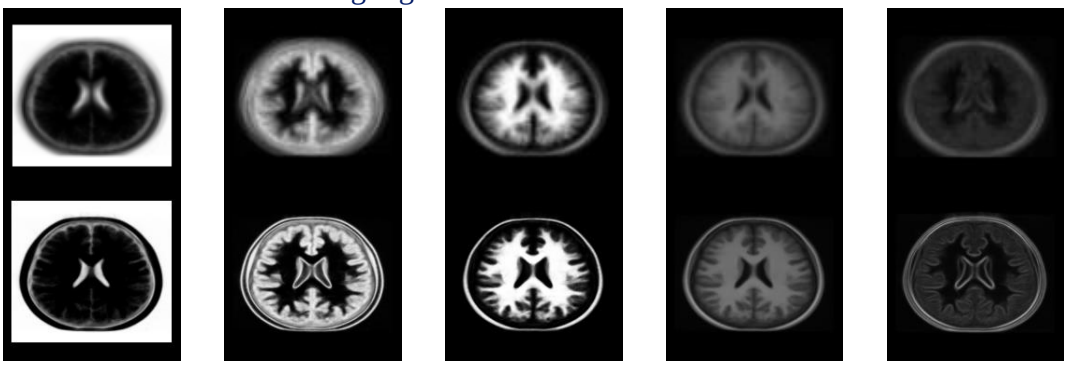


Figure 3: The top and bottom rows show initial and final component reference images respectively for the full ensemble registration (FIG). From Left to Right, the components are CSF & background, GM, and WM, intensity, and intensity gradient magnitude.

# 4 Discussion and Conclusions

We have demonstrated that integrating segmentation and registration can improve the registration results over using intensity images alone. As known correspondences achieved by registration can help segmentation to find the structures, on the other hand the found structures from segmentation can speed up the registration. Hence their integration can benefit each other.

In order to achieve our integration, we had to improve the performance of the segmentation stage. This was done by using a 2D intensity value-and-gradient histogram rather than a simple 1D intensity histogram. This gave a more robust segmentation of the images into tissue-types as shown in Figure 3.

We have also shown that adding further structural information, in the form of tissue fraction images, further improved registration results. We have shown that Ensemble registration using combinations of more than one type of images (intensity, fraction, and intensity gradient) can also produce slightly better results than previous algorithms [7]. This previous work also used segmentation but in a rather different way, in that it was used to reconstruct an artificial intensity reference image from reference fraction images.

Finally, it should be noted that whilst our results clearly show that combinations of the various types of images are needed to get the best Ensemble result, these results only used a naïve equal-weighting of components (1), and it is possible that tuning the weighting of the components could give a further improvement.

# 5 Future Work

In order to enable a direct comparison with previous work, we have reported extensive experiments on the set of 37 MR brain image slices as used previously. However, we also have available 5 other similar (but larger) datasets, with ground-truth annotation, and we are in the process of performing a similar series of experiments on these other datasets. In particular, there are no theoretical reasons why the algorithm cannot be applied to 3D data.

Our current experiments just use a repeated pairwise matching to the mean method of groupwise registration. Our next extension of our method is to consider all the image set during registration, by using a fully groupwise algorithm.

Furthermore, we also intend to consider extending our method to other types of images, such as radiographs of hands. This will be more challenging, since the tissue segmentation needed will be different. Such radiographs may introduce other problems, since they are created from projections of 3D objects, unlike MR images which are truly volumetric.

**Acknowledgment**: Ms Purwani would like to thank the Directorate General of Higher Education of Indonesia (DIKTI) for providing funding.

# References

[1] J. Ashburner and K. J. Friston. Unified segmentation. *NeuroImage*, 26: 839-851, 2005.

[2] P. A. Bromiley and N. A. Thacker, Multi-dimensional medical image segmentation with partial volume and gradient modelling, Annals of the BMVA 2008, vol. 2, 2008, pp. 1–22.

[3] W. Crum, O. Camara, and D. Hill. Generalised overlap measures for evaluation and validation in medical image analysis. *IEEE Trans. Medical Imaging*, 25:1451-1461, 2006.

[4] H. D. Gage, W. E. Snyder, and P. Santago. Quantification of brain tissue through incorporation of partial volume effects. *Medical Imaging*, 1652:8496, 1992.

[5] E. Haber and J. Modersitzki, Intensity gradient based registration and fusion of multi-modal images, MICCAI 2006. Springer Berlin Heidelberg, 2006. 726-733.

[6] E. Konukoglu, A. Criminisi, S. Pathak, D. Robertson, S. White, and K. Siddiqui, Robust linear registration of CT images using random regression forests, *SPIE Medical Imaging*, vol. 7962, 2011, p. 79621X.

[7] V. Petrovic, T. F. Cootes, C. J. Twining, and C. J. Taylor, Simultaneous registration, segmentation, and modelling of structure in groups of medical images. In *4th IEEE International Symposium on Biomedical Imaging (ISBI), From Nano to Macro*, pages 1–4, 2007.

# Automatic glottis segmentation from laryngeal high-speed videos using 3D active contours

Fabian Schenk[1]
fabian.schenk@student.tugraz.at
Martin Urschler[2]
martin.urschler@cfi.lbg.ac.at
Christoph Aigner[3]
christoph.aigner@student.tugraz.at
Imme Roesner & Philipp Aichinger[4]
philipp.aichinger@meduniwien.ac.at
Horst Bischof[1]
bischof@icg.tugraz.at

[1] Inst. f. Computer Graphics & Vision
Graz University of Technology, Austria
[2] Ludwig Boltzmann Inst. f. Clinical
Forensic Imaging, Graz, Austria
[3] Signal Processing & Speech
Communication Laboratory
Graz University of Technology, Austria
[4] Dept. of Otorhinolaryngology
Div. of Phoniatrics-Logopedics
Medical University Vienna, Austria

**Abstract**

Laryngeal high-speed videos are a state of the art method to investigate vocal fold vibration but the vast amount of data produced prevents it from being used in clinical applications. Segmentation of the glottal gap is important for excluding irrelevant data from video frames for subsequent analysis. We present a novel, fully automatic segmentation method involving rigid motion compensation, saliency detection and 3D geodesic active contours. By using the whole color information and establishing spatio-temporal volumes, our method deals with problems due to low contrast or multiple opening areas. Efficient computation is achieved by parallelized implementation using modern graphics adapters and NVidia CUDA. A comparison to a semi-automatic seeded region growing method shows that we achieve improved segmentation accuracy.

## 1 Introduction

In our service oriented society speech is the main form of communication between people and has gained a tremendous economic value. In the US around 60% of the jobs require communication skills [8] and therefore voice disorders are a factor in a country's economy. Diagnosis and classification of these disorders have become important research topics in recent years. In this context laryngeal high-speed videos (LHSVs) have emerged as a very sophisticated tool [3]. An important limitation of this method is the vast amount of video material produced in a single investigation, which makes it nearly impossible to use in clinical applications. Therefore, methods for automatically processing LHSVs with the aim of detecting vocal fold vibration patterns are highly relevant for voice disorder research. Especially the automated segmentation of the glottis area, which is the opening between the vocal

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

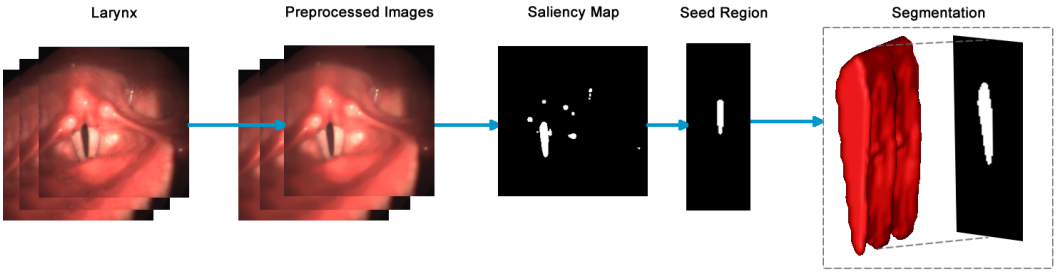



Figure 1: Image processing pipeline for automatic glottis segmentation.

folds (see Fig. 1), is an important preliminary step for later assessment of spatio-temporal plots (i.e., phonovibrograms [7]). Typical obstacles for the segmentation are the drift of the glottis due to breathing and patient movement, fluid artifacts, or brightness changes and contrast inadequacies during acquisition. The clinical standard method [6] is far from ideal in terms of accuracy and user input efficiency. In this work we propose a method to automatically segment the glottis from LHSVs using the full color information, taking motion compensation into account and processing the 3D temporal volume at once. Efficiency is achieved by a parallelized implementation using modern graphics adapters.

## 1.1 Related work

One of the first LHSV segmentation methods was presented by Lohscheller et al. [6] based on a seeded region growing (SRG) algorithm. The manual seed point and threshold selection on a large number of frames throughout the video is a time consuming task, that crucially depends on the choice of a proper homogeneity criterion. Another drawback of this method is its lack of motion compensation. Demeyer et al. [4] extended this method to get rid of user intervention. After locating the maximal glottal opening in every cycle, they look for dark elliptical regions to estimate glottis center and area. The center point is used as seed in the following SRG algorithm, which continuously adapts its threshold until the segmented region is larger than the previously estimated one. This segmentation result is adapted to the other frames in the cycle by using level sets. Similar to [6] this algorithm has problems due to leaking of the SRG and the parameters of the 2D level set are crucial but hard to select. Level set propagation heavily depends on the SRG initialization. Karakozoglou et al. [5] proposed a way to localize the glottis and find frames with the maximum glottal opening to calculate a bounding box for every cycle. The following image segmentation uses frame by frame 2D active contours. For this method the starting curve is of great importance and they propose two automatic ways for curve initialization on the landmark frames. Despite showing promising results, initital curve calculation per frame is not very robust due to leaking, leading to the need for manual post-correction.

Our contribution is an algorithm that extends these methods by performing a 3D segmentation on the spatio-temporal volumes at once, making use of the full color information. We incorporate a motion compensation step to correct for rigid patient or camera drift in a pre-processing step and tackle the leaking problem by a salient region detection initializing the segmentation. No user intervention or manual refinement is required in our method.

# 2 Method

We set up an image processing pipeline as depicted in Fig. 1 to tackle the specific problem of glottis segmentation from LHSVs. The typical behavior of the glottis during voice production is a repeated opening and closing process. The frame rate is much higher than the fundamental frequency of the vocal folds, which allows treating the glottography videos as 3D temporal volumes. To get rid of artifacts we apply an edge-preserving denoising filter to the 3D temporal volume, followed by a rigid registration of subsequent frames. The core of our method is the saliency detection using a boolean map approach, which gives us the seed regions located in the interior part of the glottis. The following 3D segmentation uses these seeds to compute the opening and closing glottis volumes.

## 2.1 Preprocessing

Edge-preserving denoising is used on the frames to get rid of light and fluid artifacts while keeping important edge information intact. Our implementation follows the total variation based 3D denoising paradigm that regularizes the L1 norm of the gradients in the denoising solution. The continuous convex optimization scheme is based on the primal-dual algorithm from Chambolle and Pock [1]. We apply this method to each color channel separately. Preventing smoothing over the important edge information is crucial for later segmentation.

To get rid of global patient or camera movement, the frames have to be registered onto each other. The high frame rate of the recording provides a certain advantage for the image registration because we can assume that there is no movement between two or even more consecutive frames, since camera or patient drifts occur more slowly over the videos. Therefore, to speed up this process we only register every tenth image to our fixed image and apply the transformation to the four frames before and the five after the registered frame. For image registration we use a slightly modified version of the *magnitude difference minimization method* proposed by Delisyki et al. [2] by additionally taking rotation into account. This is essentially an intra-modality registration using the sum of absolute color pixel differences as a similarity measure. Since individual differences between frames are small, it is feasible to solve this registration problem globally by an exhaustive search strategy over the two translation and the rotation parameters in a limited parameter range. Efficient computation of this step is achieved by a parallel NVidia CUDA GPU implementation.

## 2.2 Saliency detection

The saliency detection is crucial in our image processing pipeline as it provides seed points in the interior of the glottis that initialize the following segmentation algorithm. We face a saliency detection problem, where we have to calculate a saliency map from an RGB-image to represent the interesting regions, defined as areas fully surrounded by background. The glottal opening is a dark area surrounded by tissue, thus fulfilling these requirements. Without the previous denoising step fluid artifacts or tissue anomalies would be detected as interesting regions as they are usually areas on a homogeneous background.

### 2.2.1 Boolean Map based Saliency

We adapted a method proposed by Zhang and Sclaroff [10], as depicted in Fig. 2. The main idea is to generate boolean maps by thresholding the different color channels. We use the

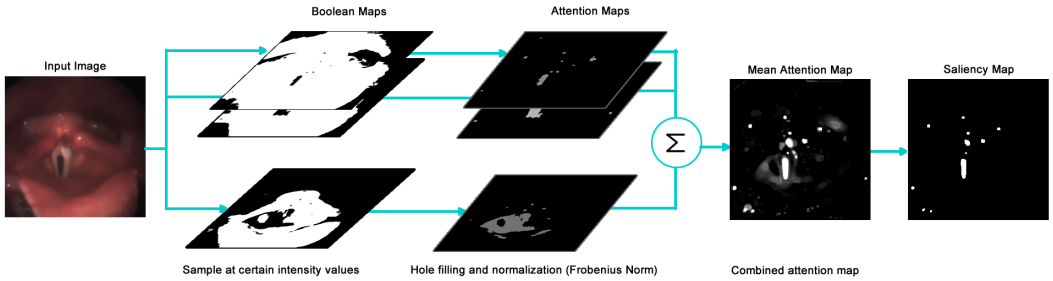



Figure 2: Saliency detection based on boolean maps, adapted from [10].

CIE Lab color space due to its perceptual uniformity with all channels scaled to [0,255]. In order to generate boolean maps we iterate through all the channels and sample them at a certain intensity threshold value $\theta$, which is increased by a fixed step size $\delta$ after each iteration. An attention map is calculated by closing the surrounded regions or holes of the boolean map followed by a division by its Frobenius norm to emphasize small concentrated areas. The results are then combined to a mean attention map and a threshold operation filters out small and irrelevant signals. With the use of a connected component analysis the darkest and biggest area is found and used as seed region for the segmentation.

### 2.3 Image Segmentation

The 3D spatio-temporal volume segmentation problem can be formulated as a Markov Random Field and solved using graph cuts. We formulate this problem as the equivalent continuous formulation using the calculus of variation, which leads to a convex geodesic active contour based on the weighted total variation (GAC-TV). Minimization of this functional has been demonstrated in [9] and leads to a temporally smooth globally optimal minimal surface solution. We work solely on the gradient information of the 3D temporal volume, computed from the RGB values according to $V_{grey} = 0.2126R + 0.7152G + 0.0722B$.

The initialization regions come from the salient region detection step described above and the 3D segmentation iteratively evolves them towards the edges of the temporal volume. This segmentation method in 3D gives an advantage over traditional active contour algorithms like [4, 5], since it allows to fill in topologically disconnected regions in single frames.

## 3 Experiments and Results

Our algorithm is intended to process LHSVs that typically consist of 4000 frames/s at a resolution of $256 \times 256$ pixels covering two seconds of phonation. To quantitatively evaluate our method, on the one hand temporal sub-volumes of consecutive frames and on the other hand randomly chosen single frames from a number of videos were selected. Our experimental setup consists of two 60- and one 30-frame video sequences as well as 25 single frames randomly picked from different videos, involving healthy and disordered subjects. For all frames an experienced computer vision researcher has performed a manual segmentation of the glottis, which was investigated and corrected by an expert in kymography and voice disorders. These segmentations are the ground truth to which we compare our proposed algorithm as well as an implementation of the clinical standard [6]. For quantitative

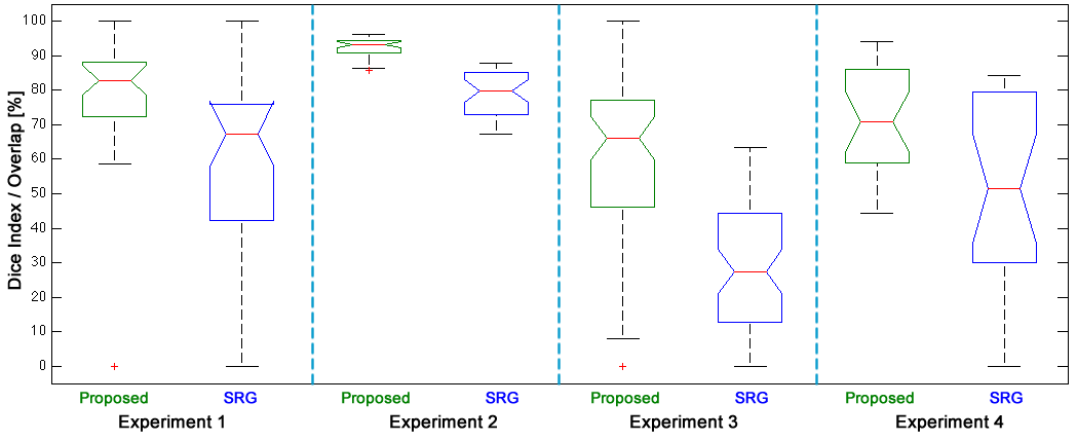



Figure 3: Box-whisker plots showing four different experiments with the proposed method in green and SRG [6] in blue. Exp. 1 and 2 are the two 60-frame videos, Exp. 3 is the 30-frame sequence and Exp. 4 represents the 25 randomly picked frames.

evaluation we compute the Dice coefficient as a segmentation overlap measure for each segmentation approach and the ground truth segmentation. The Dice coefficient (DC) is defined as $DC = \frac{2|A \cap B|}{|A|+|B|}$, with $A$ the ground truth and $B$ a segmentation. It computes a value between 0 and 1, which we present as an overlap percentage. All the calculations were performed on a Linux machine using an Intel Core i7 and an NVidia GeForce GTX 580 graphics adapter.

The segmentation results can be seen in Fig. 3, where the proposed method is far more accurate with median values of 82.9, 93.2, 66.1 and 70.1 % compared to 67.4, 79.8, 27.5 and 51.5 % of the SRG method. Some of the outliers can be explained due to the nature of the dice index, which over-emphasizes relatively small mistakes for small structures (i.e. missing a ground truth segmentation that only consists of a single pixel would result in overlap $DC = 0$). There is also a great advantage when it comes to speed, as the proposed fully automatic method takes only 50 seconds to segment 100 frames compared to 3.5 minutes of the SRG method, which also highly depends on the experience of the user. Figure 4 shows exemplary qualitative segmentation results of the two methods in the three different videos. The low contrast at the bottom of Fig. 4(a,i) poses a leaking problem for the threshold based SRG method (see Fig. 4(c,k)), whereas the proposed method does not have this problem (see Fig. 4(b,j)). Figure 4(j) indicates a problem of the proposed algorithm. Due to the narrow connection between the bottom and top region, the edges in this part of the image are not sufficient to detect, which makes it expensive for the GAC-TV segmentation to connect these two areas. The frames in Fig. 4(e-h) show that on an image with a high contrast difference and a big opening area both methods work well.

# 4 Conclusion

A fully automatic algorithm for glottis segmentation from LHSVs has been proposed, which combines motion compensation, robustness to video artifacts and a fully 3D segmentation algorithm to efficiently process glottography video material. A comparison with the clinical standard method has revealed a higher segmentation accuracy and at the same time requires no user interaction, thus our expected run-time scales much better to the real-world applica-

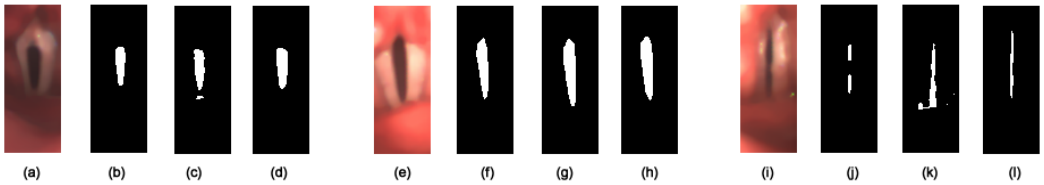



Figure 4: Qualitative results. (a,e,i) original frames, (b,f,j) segmentation results of the proposed method (c,g,k), segmentation results of the SRG method and (d,h,l) ground truth.

tion of segmenting thousands of video frames. Future work will concentrate on evaluating a larger database as well as a comparison to further related algorithms.

# References

[1] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal Math Imaging and Vision*, 40(1):120–145, 2011.

[2] D. Deliyski, S. Cieciwa, and T. Zielinski. Fast and robust endoscopic motion estimation in high-speed lyrngoscopy. In *Proc Conf Advances in Quantitative Laryngology, Voice and Speech Research (AQL)*, 2006.

[3] D. Deliyski, P. Petrushev, H. Bonilha, T. Gerlach, B. Martin-Harris, and R. Hillman. Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and evolution. *Folia Phoniatrica et Logopaedica*, 60:33–44, 2008.

[4] J. Demeyer, T. Dubuison, B. Gosselin, and M. Remacle. Glottis segmentation with a high-speed glottography: a fully automatic method. In *3rd Int. Workshop Adv. Voice Funct. Assess.*, 2009.

[5] S. Karakozoglou, N. Henrich, C. d'Alessandro, and Y. Stylianou. Automatic glottal segmentation using local-based active contours and application to glottovibrography. *Speech Communication*, 54(5):641–654, 2012.

[6] J. Lohscheller, H. Toy, F. Rosanowski, U. Eysholdt, and M. Doellinger. Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos. *Medical Image Analysis*, 11(4):400–413, 2007.

[7] J. Lohscheller, U. Eysholdt, H. Toy, and M. Doellinger. Phonovibrography: Mapping high-speed movies of vocal fold vibrations into 2D diagrams for visualizing and analyzing the underlying laryngeal dynamics. *IEEE Trans Med Imag*, 27:300–309, 2008.

[8] R.J. Ruben. Redefining the survival of the fittest: communication disorders in the 21st century. *Laryngoscope*, 110:241–245, 2000.

[9] M. Unger, T. Pock, W. Trobin, D. Cremers, and H. Bischof. TVSeg – interactive total variation based image segmentation. In *Proc British Machine Vision Conf*, 2008.

[10] J. Zhang and S. Sclaroff. Saliency Detection: A Boolean Map Approach. In *Proc Int Conf Computer Vision (ICCV)*, pages 153–160, 2013.

# Regularized Geometric Hulls for Nucleus Segmentation in Microscopy Images

Marco Körner
marco.koerner@uni-jena.de
Mahesh V. Krishna
mahesh.v.krishna@uni-jena.de
Herbert Süße
herbert.suesse@uni-jena.de
Wolfgang Ortmann
wolfgang.ortmann@uni-jena.de
Joachim Denzler
joachim.denzler@uni-jena.de

Computer Vision Group
Friedrich Schiller University of Jena
Jena, Germany
http://www.inf-cv.uni-jena.de

**Abstract**

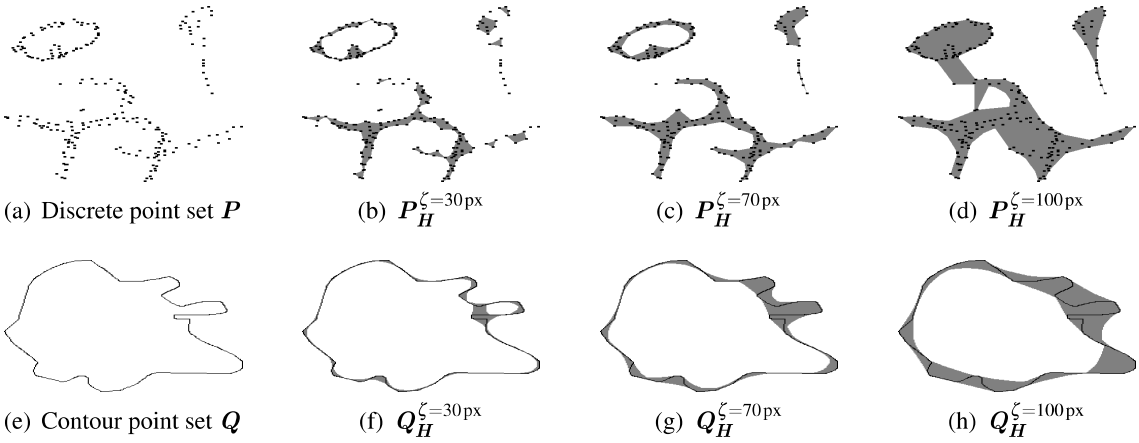
As segmentation of cells from microscopy images is a very complex and error-prone task, we present a new unsupervised and fully-automatic method based on *Regularized Geometric Hulls (RGH)*. While it performs up to two orders of magnitudes faster than state-of-the-art methods, wide-range experiments performed on the recent *CellCognition* dataset demonstrate its robustness and accuracy.

## 1 Introduction

During the decades of research in the field of microscopy-based tissue analysis, the identification and segmentation of individual cells is one of, if not the most important and fundamental step within the whole processing pipeline. While classically this task had to be performed manually, more and more automatic or semi-automatic methods came into the focus of modern computer vision research. As this problem is highly affected by image degradations such as blur or pixel noise, traditional contour segmentation methods are likely to fail in these scenario. Furthermore, other approaches require manual annotation of cells for localization or employ computationally expensive machine learning techniques for building complex models of cell appearance. Segmentation of cells from images has been tackled with a wide variety of tools, ranging from simple thresholding to machine learning-based methods [5]. Many of them tackle specific types of microscopy images and are application specific.

For instance, approaches based on *Seeded Watershed Transform* are well-established for cell segmentation [6]. However, the quality of the results depend heavily on the initial seeds and the approach often produces over-segmented images. These shortcomings necessitate further processing steps, which result in computationally complex algorithms, and a different approach is called for. Other popular approaches use machine learning methods. Arteta *et al.* [1] employ a SVM classifier for categorizing each pixel to belong to either cell foreground or

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.



(a) Discrete point set $\boldsymbol{P}$ (b) $\boldsymbol{P}_H^{\zeta=30\text{px}}$ (c) $\boldsymbol{P}_H^{\zeta=70\text{px}}$ (d) $\boldsymbol{P}_H^{\zeta=100\text{px}}$

(e) Contour point set $\boldsymbol{Q}$ (f) $\boldsymbol{Q}_H^{\zeta=30\text{px}}$ (g) $\boldsymbol{Q}_H^{\zeta=70\text{px}}$ (h) $\boldsymbol{Q}_H^{\zeta=100\text{px}}$

Figure 1: Influence of different values for the regularization parameter $\zeta$ on geometric hull construction for the case of (a)–(d) dense point sets and (e)–(h) contour point sets. Note that the RGH preserves inner contours.

background. This binarization is further analyzed for connected components to achieve final segmentation results. While this approach shows promising results, pixel-wise operations on high-dimensional image sequences common in time-lapse microscopy videos make it unsuitable for fast processing. Furthermore, it depends on the availability of manually labeled training data, which is often difficult or impossible to supply.

In contrast, we propose a method for fast and robust joint cell detection and segmentation based on *Regularized Geometric Hulls (RGH)* [7] which is designed to be unsupervised, acting fully automatic, and requires only one rather simple parameter to be adjusted.

## 2 Regularized Geometric Hull

This section is dedicated to briefly introduce into the concept of *Regularized Geometric Hulls (RGH)* recently presented by Süße *et al.* [7], which is the core ingredient of our proposed method. These can be regarded as an extension of classical convex hulls.

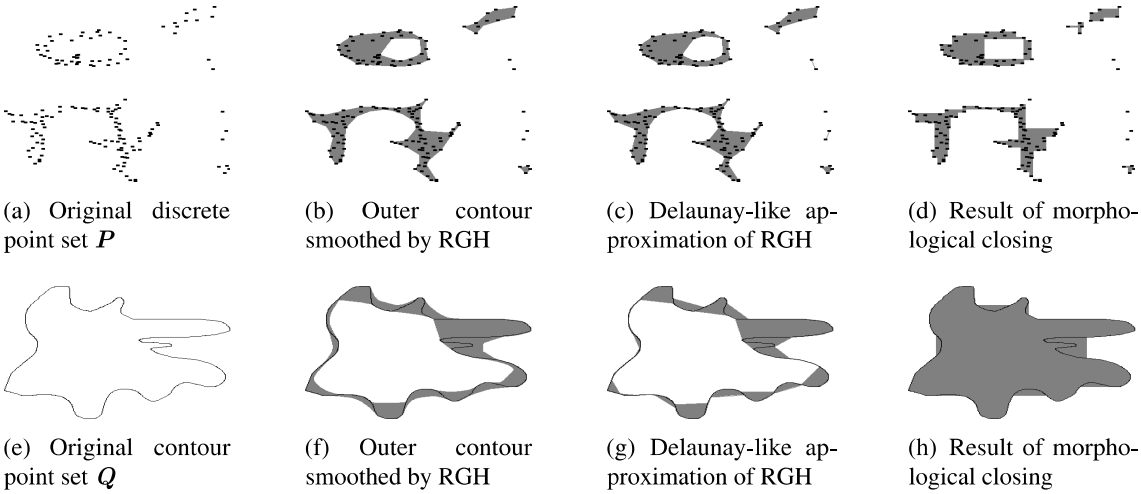
**Definition 1** (Convex hull)**.**
*Given an arbitrary set of Euclidean points* $\boldsymbol{P} = \{\boldsymbol{p} | \boldsymbol{p} \in \mathbb{R}^2\}$ *(or discrete lattice points* $\boldsymbol{Q} = \{\boldsymbol{q} | \boldsymbol{q} \in \mathbb{Z}^2\}$*), the* Convex Hull $\boldsymbol{P}_H \supseteq \boldsymbol{P}$ *(*$\boldsymbol{Q}_H \supseteq \boldsymbol{Q}$*) is the smallest convex subset of* $\mathbb{R}^2$ *that contains* $\boldsymbol{P}$ *(*$\boldsymbol{Q}$*).*

Various algorithms for the computation of the convex hull with fair runtime complexity $\mathcal{O}(n \log n)$ exist in the field of computational geometry. As this convex hull approximation of points is insufficiently coarse for most computer vision problems, we aim to allow a certain degree of non-convexity—*i.e.* concavity—while computing geometric hulls.

**Definition 2** (adjacent points)**.**
*Let* $d : \boldsymbol{M}^2 \mapsto \mathbb{R}$ *be a metric in* $\boldsymbol{M}^2$*. Two points* $\boldsymbol{p}, \boldsymbol{q} \in \boldsymbol{M}^2$ *are called* $\zeta$-adjacent *wrt. a non-negative constant* $0 \leq \zeta \in \mathbb{R}$*, if*

$$\boldsymbol{p} \sim^{\zeta} \boldsymbol{q} \Leftrightarrow d(\boldsymbol{p}, \boldsymbol{q}) \leq \zeta \quad . \tag{1}$$

**Definition 3** (Regularized Geometric Hull (RGH) for Euclidean point sets)**.**
*Let* $\boldsymbol{P} = \{\boldsymbol{p} | \boldsymbol{p} \in \mathbb{R}^2\}$ *be an arbitrary set of Euclidean points and* $\Delta(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) \subseteq \boldsymbol{P}$ *the set of all points* $\boldsymbol{p}_i \in \boldsymbol{P}$ *enclosed by the triangle defined by the 3-tuple* $(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) \in \boldsymbol{P}^3$*.*



(a) Original discrete point set $\boldsymbol{P}$ (b) Outer contour smoothed by RGH (c) Delaunay-like approximation of RGH (d) Result of morphological closing

(e) Original contour point set $\boldsymbol{Q}$ (f) Outer contour smoothed by RGH (g) Delaunay-like approximation of RGH (h) Result of morphological closing

Figure 2: The *RGH* can be approximated by using Delaunay-like triangulations of the input point cloud $\boldsymbol{P}$ or contour $\boldsymbol{Q}$, which realizes a morphological closing operation with adaptive, polygonial structure elements.

*Further, let*

$$\boldsymbol{P}_{\Delta}^{\zeta} = \left\{ \Delta(\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3) \in \boldsymbol{P}^3 \mid \boldsymbol{p}_1 \sim^{\zeta} \boldsymbol{p}_2 \wedge \boldsymbol{p}_2 \sim^{\zeta} \boldsymbol{p}_3 \wedge \boldsymbol{p}_1 \sim^{\zeta} \boldsymbol{p}_3 \right\}, \quad 0 \leq \zeta \in \mathbb{R}, \tag{2}$$

*be the set of all triangles of $\zeta$-adjacent points in $\boldsymbol{P}$. Then, the union set*

$$\boldsymbol{P}_{H}^{\zeta} = \bigcup \boldsymbol{P}_{\Delta}^{\zeta} \cup \boldsymbol{P} \tag{3}$$

*is called the* Regularized Geometric Hull (RGH) *of $\boldsymbol{P}$ wrt. $\zeta$.*

**Definition 4** (Regularized Geometric Hull (RGH) for discrete lattice point sets)**.**
*Let $\boldsymbol{Q} = \left\{ \boldsymbol{q} | \boldsymbol{q} \in \mathbb{Z}^2 \right\}$ be an arbitrary set of discrete lattice points and $\Delta(q_1, q_2, q_3) \subseteq \boldsymbol{Q}$ the set of all triangular points $\boldsymbol{q}_i \in \boldsymbol{Q}$ induced by the lattice points $(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) \in \boldsymbol{Q}^3$. Let further be*

$$\boldsymbol{Q}_{\Delta}^{\zeta} = \left\{ \Delta(\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3) \in \boldsymbol{Q}^3 \mid \boldsymbol{q}_1 \sim^{\zeta} \boldsymbol{q}_2 \wedge \boldsymbol{q}_2 \sim^{\zeta} \boldsymbol{q}_3 \wedge \boldsymbol{q}_1 \sim^{\zeta} \boldsymbol{q}_3 \right\}, \quad 0 \leq \zeta \in \mathbb{R}, \tag{4}$$

*the set of all* non-degenerated *triangles of $\zeta$-adjacent points in $\boldsymbol{Q}$. Then, the union set*

$$\boldsymbol{Q}_{H}^{\zeta} = \bigcup \boldsymbol{Q}_{\Delta}^{\zeta} \cup \boldsymbol{Q} \tag{5}$$

*is called the* Regularized Geometric Hull (RGH) *of $\boldsymbol{Q}$ wrt. $\zeta$.*

Clearly, the properties of the geometric hull $\boldsymbol{P}_{H}^{\zeta}$ strictly depend on the choice of the structure parameter $\zeta$, as can be seen in Fig. 1. If $\zeta = 0$, the geometric hull $\boldsymbol{P}_{H}^{0}$ is identical to the point set $\boldsymbol{P}$. In turn, if $\zeta \to \infty$, the geometric hull $\boldsymbol{P}_{H}^{\zeta}$ converges to the convex hull $\boldsymbol{P}_{H}$ yielding the relation

$$\boldsymbol{P} = \boldsymbol{P}_{H}^{0} \subseteq \boldsymbol{P}_{H}^{\zeta} \subseteq \boldsymbol{P}_{H}^{\infty} = \boldsymbol{P}_{H}, \quad 0 \leq \zeta \in \mathbb{R}. \tag{6}$$

The structure parameter $\zeta$ regularizes the convexity or concavity of the geometric hull. In contrast to the classical convex hull $\boldsymbol{P}_{H}$, the RGH $\boldsymbol{P}_{H}^{\zeta}$ describes the geometric structure of the input data $\boldsymbol{P}$ by a set of outer and inner contours.

## 2.1 Approximation

Def. 3 and 4 suggest that the computation of the RGH shows cubic complexity $\mathcal{O}(n^3)$ in the number $n$ of input points which makes its application unhandy in realistic scenarios. For this

reason, we approximate the underlying set of triangles $\boldsymbol{P}_\Delta^\zeta \subseteq \boldsymbol{P}^3$ by a modification of the *Delaunay triangulation* [2] $\boldsymbol{P}_\Delta^{\zeta\mathrm{DT}} \subseteq \boldsymbol{P}^3$, which is supposed to be a triangular tessellation of a given point set $\boldsymbol{P}$ optimal wrt. a given distance criterion. All triangles obtained by Delaunay's triangulation algorithm are drawn into a binary image, wherefrom final contours are extracted afterwards. Hence, the Delaunay triangulation $\boldsymbol{P}_\Delta^{\zeta\mathrm{DT}}$—restricted to triangles with edge lengths smaller than $0 \leq \zeta \in \mathbb{R}$—can be interpreted as an approximation of the previously used set $\boldsymbol{P}_\Delta^\zeta$ of triangles of $\zeta$-adjacent points. The images shown in Fig. 2(c) and 2(g) show the results of the RGH algorithm using the Delaunay-like approximation for both a discrete point set and a set of discrete contour points, respectively. Compared to the results of the original RGH algorithm shown in Fig. 2(b) and 2(f), the results are reasonably good and useful for further processing. As the Delaunay algorithm shows solely quasi-logarithmic complexity $\mathcal{O}(n \log n)$ in the number $n$ of vertices, this approximation remarkably speeds up the whole computation and allows for real-time performance.

## 2.2 Geometric Interpretation

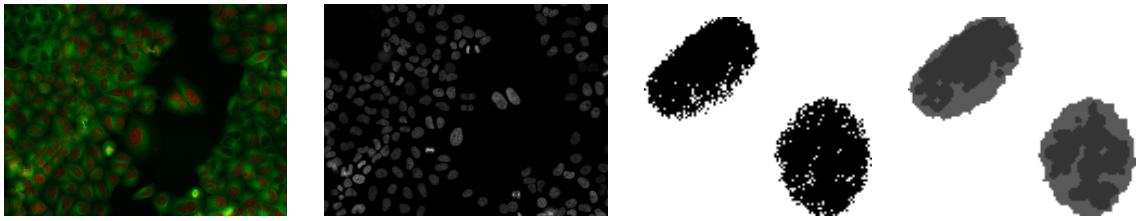
When analyzing the results of the RGH or their approximation based on Delaunay triangulation, one can observe various handy properties useful for further processing of precomputed contours or point sets. First, the RGH augments the input contour or dense point set $\boldsymbol{P}$ by a *geometric orientation*, enabling us to easily apply neighborhood-based methods for shape recognition. While doing so, in the case of discrete point sets and dependent from the structure parameter $\zeta$, existing holes are preserved and isolated objects are detected, as can be seen in Fig. 2(b) and (c).

Second and as illustrated in Fig. 2, the presented method performs a *smoothing* of a given contour which can be used to deal with inaccurate segmentation. Again, the impact of this smoothing is tuned by the structure parameter $\zeta$. In terms of computational geometry, this is comparable to the *morphological closing* operator applied to binary objects. So, RGHs can be characterized as such a closing operation based on adaptive, polygonial structure elements. In contrast to classical morphological closing with constant regular structure elements (*e.g.* rectangles, ellipses), the results of RGHs constructed from contours or dense point sets appear much more smooth and intuitive, as can be seen in Fig. Fig. 2(d) and (h), respectively.

# 3 Evaluation

## 3.1 Dataset

As annotated image data of cells is very rare, most of the existing approaches were evaluated on rather small non-public databases. For instance, Arteta *et al.* [1] used three datasets of 12 to 22 images in their experiments. In contrast, the *CellCognition* dataset presented by Held *et al.* [4] consists of 7 sequences of 206 frames of size $1392 \times 1040$ px, each showing RNAi treated human HeLa Kyoto cells expressing fluorescent H2B-mCherry (orange) and $\alpha$-tubulin (green). In total, the dataset contains 363,120 annotated cell objects with automatically generated ground truth data providing centroids and bounding boxes. Fig. 3 shows an example frame of the data-set, with the full recording and the segregated nuclei. Apart from the huge number of objects, the data-set is made more challenging by low contrast in many cases, and overlapping cells due to limits in the thinness of slide preparations.



(a) microscopic image (b) corresponding nuclei expressing H2B-mCherry (c) Thresholded nuclei (d) RGH result

Figure 3: Images from the CellCognition dataset (a) are initially thresholded (c) and further used to extract nuclear boundaries by RGH (d).

### 3.2 Experimental Setup

As the proposed method is designed in a very general way, the overall framework is rather straightforward. After binary thresholding the input images (*cf*. Fig. 3(c)), they are directly passed to the RGH module to create foreground-background images as exemplary shown in Fig. 3(d). The values of the structure parameter $\zeta$ are varied in order to show its influence on the overall results. Cells are finally extracted by subsequent contour extraction.

For comparison to other methods we used two popular quantification methods. First, suggested by Arteta *et al*. [1], we assumed each cell detected by our system to be a true positive, if the distance between its centroid and the centroid of the nearest ground truth cell is smaller than a predefined threshold $\rho_{\text{dist}}$. Usually, this threeshold is set to the radius of the smallest cell expected to appear in the data. Second and more commonly used in object detection, a detected cell is assumed to match a ground truth cell if the ratio

$$r\left(B_{\text{det}}, B_{\text{gt}}\right) = \frac{\text{area}\left(B_{\text{det}} \cap B_{\text{gt}}\right)}{\text{area}\left(B_{\text{det}} \cup B_{\text{gt}}\right)} \tag{7}$$

of the union and intersection areas of respective bounding boxes $B_{\text{det}}$ and $B_{\text{gt}}$ exceeds the value of $\rho_{\text{area}} = 0.5$ (*cf*. *Pascal Criterion* [3]).

### 3.3 Results

In order to evaluate the performance of our methods, we compared to the state-of-the-art method presented by Arteta *et al*. [1]. From the numbers reported in Tab. 1, it can be observed that our approach performs better or at least comparable to the competing method. Intuitively, the structure parameter $\zeta$ only slightly influences the overall accuracy. Wrong segmentation mainly occurs when two objects are closer than the structure parameter $\zeta$. Furthermore it has to be noted that the evaluation based on cell centroid distances yields higher accuracy. This should be taken into account for further comparisons. The most remarkable observation arises from the computational analysis. While the supervised method of Arteta *et al*. [1] requires a very time-consuming training phase, even the testing step is rather slow. In contrast, our approach performs localization and segmentation almost two orders of magnitudes faster, which is a great advantage in real-world applications.

## 4 Summary

We proposed a method for detecting and segmenting cell in microscopy images relying on *Regularized Geometric Hulls (RGH)* only incorporating one single parameter which is easy to

| Method | Accuracy | | Runtime | |
|---|---|---|---|---|
| | Centroid distance $\rho_{\text{dist}} = 20\,\text{px}$ | Pascal criterion $\rho_{\text{area}} = 0.5$ | training | testing |
| our approach | | | — | 1.96 s/frame |
| $\zeta = 3\,\text{px}$ | 91.65% | 85.75% | | |
| $\zeta = 4\,\text{px}$ | 91.75% | 85.01% | | |
| $\zeta = 5\,\text{px}$ | 92.11% | 85.77% | | |
| Arteta *et al.* [1] | 88.79% | 86.20% | 6231.59 s | 130.66 s/frame |

Table 1: Evaluation of our approach and comparison to state-of-the-art methods.

tune. In contrast to competing approaches, our method does not require any supervision, prior learning phases or manual annotation. Extensive experiments on the recent *CellCognition* dataset verified the accuracy of our approach and demonstrated its superior runtime behavior.

Subjected to future work, we intend to predict the structure parameter $\zeta$ adaptively while segmentation or learn it from training data and to use more sophisticated preprocessing of the input data compared to binary thresholding as used in our framework.

## Acknowledgements

The authors thank Christian Wojek and Stefan Saur for useful discussions and suggestions. Mahesh Venkata Krishna is supported by Carl Zeiss AG through the “Pro-Excellenz” scholarship of the Federal State of Thuringia, Germany.

## References

[1] C. Arteta, V. Lempitsky, J. Noble, and A. Zisserman. Learning to detect cells using non-overlapping extremal regions. In *MICCAI*, pages 348–356. Springer-Verlag, 2012.

[2] B. Delaunay. Sur la sphère vide. *USSR Academy of Sciences*, (6):793–800, 1934.

[3] M. Everingham, L van Gool, C. Williams, J. Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. *IJCV*, 88(2):303–338, 2010.

[4] M. Held, M. Schmitz, B. Fischer, T. Walter, B. Neumann, M. Olma, M. Peter, J. Ellenberg, and D. Gerlich. Cellcognition: Time-resolved phenotype annotation in high-throughput live cell imaging. *Nature Methods*, 7(9):747–754, 2010.

[5] E. Meijering. Cell segmentation: 50 years down the road. *Signal Processing Magazine, IEEE*, 29(5):140–145, Sept 2012.

[6] F. Meyer and S. Beucher. Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1(1):21–46, 1990.

[7] H. Süße, W. Ortmann, J. Lautenschläger, C. Lautenschläger, M. Körner, J. Großkreutz, and J. Denzler. Quantitative analysis of pathological mitochondrial morphology in neuronal cells in confocal laser scanning microscopy images. In *IWBBIO*, pages 1290–1301, 2014.

# Feature Analysis and Classification

# Recognizing cardiac magnetic resonance acquisition planes

Jan Margeta[1]
jan@kardio.me
Antonio Criminisi[2]
antcrim@microsoft.com
Daniel C Lee[3]
dlee@northwestern.edu
Nicholas Ayache[1]
nicholas.ayache@inria.fr

[1] Asclepios, Inria Sophia Antipolis,
2004 Route des Lucioles
Sophia Antipolis, France

[2] Machine learning and perception group
Microsoft Research,
21 Station Rd, Cambridge, UK

[3] Feinberg Cardiovascular Research
Institute, Northwestern University,
Feinberg School of Medicine,
Chicago, USA

**Abstract**

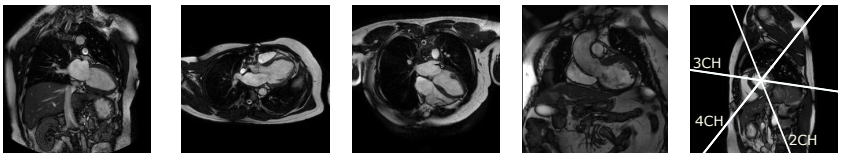
In this paper we propose a method for automatic wrangling of missing or noisy acquisition plane information of cardiac magnetic resonance images in order to simplify case filtering and image lookup in large collections of cardiac data. To recognize standard cardiac planes we use features based on image miniatures combined with a decision forest classifier. We show that augmenting the dataset with a set of nondestructive transformations can improve classification accuracy. Our approach compares favorably to the state of the art while requiring fewer manual annotations.

## 1 Introduction

Cardiac magnetic resonance image acquisitions are mostly performed along several oblique axes with respect to the structures of the heart instead of traditional body planes (coronal, axial and sagital) to better visualize and quantify different aspects of the heart. Knowledge of these planes is also essential for data collection organization, selecting adapted image processing algorithms, grouping of related slices into volumetric image stacks, filtering of cases based on completeness for a clinical study and to respond with the most relevant acquisition plane in content based image retrieval.

This orientation information is often in some way encoded within the DICOM image tags: Series Description (0008,103E) and Protocol Name (0018,1030). The way to encode this view information is however not standardized, operator errors are present or such information is often completely missing. Searching through large databases to handpick desired views from cases in the collection is therefore tedious. Previous work in this field has concentrated mainly on real-time recognition of cardiac planes in echography acquisitions. These methods are based on appearance models, registration or part detection and require additional information to train view, part [6] or landmark specific detectors [10]. Therefore

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.




(a) 2CH (b) 3CH (c) 4CH (d) LVOT (e) SAX

Figure 1: Examples of the main left ventricular long and short axis cardiac MR views.

any new view will require these extra annotations to be made. Otey *et al*. [5] on the other hand addressed this problem by using view labels only from gradient based image features.

### 1.1 Cardiac planes of acquisition

Optimal cardiac planes depend on global positioning of the heart in the thorax. This is more vertical in young individuals and more diaphragmatic in elderly. Imaging in standard cardiac planes ensures efficient coverage of relevant cardiac territories and enables comparisons across modalities, thus enhancing patient care and cardiovascular research. A good overview of standard cardiac acquisition planes can be found in [8].

**Left ventricular long axis acquisition planes.** These are usually acquired as 2D or cine 2D+t stack. The 2-chamber, 3-chamber, and 4-chamber views (figs. 1(a) to 1(c)) are used to visualize different regions of the left atrium, mitral valve apparatus, and left ventricle. The 3-chamber and left ventricular outflow tract (fig. 1(d)) views provide visualization of the aortic root from two orthogonal planes. The 4-chamber view enables visualization of the tricuspid valve and right atrium.

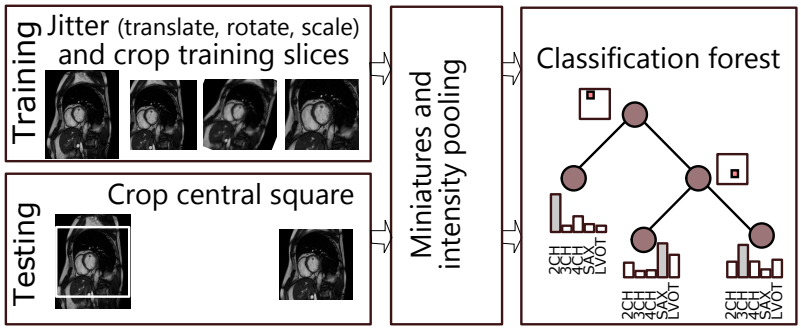
**Short axis acquisition planes.** Short axis slices (fig. 1(e)) are oriented parallel to the mitral valve ring. These are acquired regularly spaced from the cardiac base to the apex of the heart, often as a cine 3D+t stack. These views are excellent for reproducible volumetric measurements or radial cardiac motion analysis but their use is limited in atrio-ventricular interplay or valvular disease study.

## 2 Method

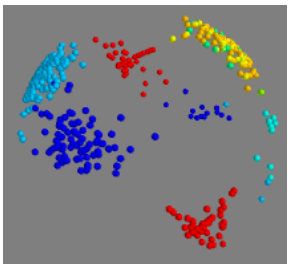
We propose automatic cardiac view recognition pipeline (Fig. 2(a)) that learns to recognize the acquisition planes directly from cardiac magnetic resonance images by combining image miniatures with classification forests. These are trained in a supervised fashion to automatically select the relevant features. To augment the training dataset we jitter the images with small translations, rotations and scales. Apart from the view annotation of the training set, no other information is necessary to classify the views.

### 2.1 Slice normals from DICOM

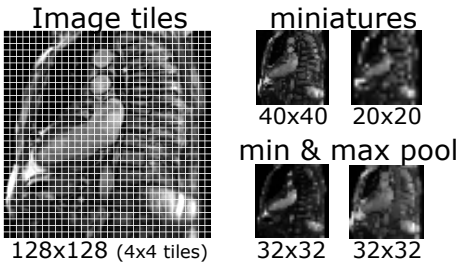
For images where the DICOM orientation (0020,0037) tag is present we can use it to predict the cardiac image acquisition plane. Similarly to [10] we extract image acquisition plane normal vectors as a cross-product of the two image orientation vectors specified in the tag and use this three-dimensional vector feature vector to describe each image (See Fig. 2(b)).




(a) Discriminative pixels from image miniatures are chosen from a random pool as features for a classification forest. We jitter the training dataset to improve robustness to differences in the acquisitions without the need for extra data.



(b) DICOM plane normals for different cardiac views.




(c) Image feature maps as downsampled images and tile intensity minima and maxima.

Figure 2: Our view recognition pipeline (above) and used features (below).

## 2.2 Image-based features

When the orientation tag is not present we rely on the image intensity information only. To capture differences between planes of acquisition, we extract a set of image based features. To allow use of this for method for all of the above mentioned applications, we use $2D$ image slices individually rather than $3D$ or $3D+t$ volumes.

**Preprocessing** Our only preprocessing steps are image central square extraction, resampling to 128x128 pixels and linear rescaling of the intensities to range between 0 and 255.

**Pooled image miniatures** Radiological images are mostly acquired with the object of interest in the center. Therefore some rough alignment of structures can be expected (See fig. 3). Image intensity samples at fixed locations can therefore provide strong clues about the position of different tissues (e.g. symmetric dark lungs or bright cavity in the center).

It has been shown before [9] that significantly down-sampled miniatures can be used

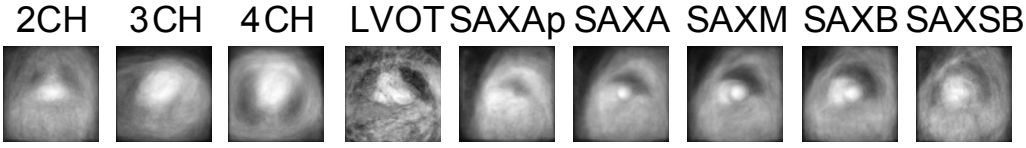



Figure 3: Image miniature averages for each cardiac view after cropping the central square. Short axis view means are displayed from the apex SAXAp to above base slices SAXSB. Distinct patterns of the bright blood pool or dark lungs can be observed for each view.

for such image recognition. In this paper we subsample the cropped centers to two fixed sizes (20x20 and 40x40 pixels). We also create another set of pooled image miniatures from each cropped center by dividing the image into non-overlapping 4x4 tiles and computing intensity minimum and maximum in each of these tiles (fig. 2(c)). As in [4], pooling can add invariance to small image translations and rotations (whose effect is within the tile size). We used directly pixel values from these miniatures sampled at random positions as features. Each image can be seen as a data point in this high-dimensional feature space.

## 2.3 Decision forest classifier

Random decision forest [1] is an ensemble classification method that consists of a set of binary decision trees. This method is computationally efficient and allows automatic selection of relevant features for the prediction. The tree structure is optimized by recursively partitioning the collection of data points $X$ into the left or right branches such that points with different labels get separated while the same label points are grouped together. At each node of the tree a feature from a randomly drawn subset of all features is chosen such that impurity $I$ in both branches is minimized. We weight samples from the under-represented classes more and we set sample weights inversely proportional to class probabilities and normalize them such that sum $w_{left} + w_{right}$ to equal to one at each node.

$$I(X, \theta) = w_{left} H(X_{left}) + w_{right} H(X_{right}) \tag{1}$$

$H$ is weighted entropy and $X_{left}$ and $X_{right}$ are point subsets falling to either the left or the right branch, based on the tested feature value and threshold. Similarly, $w_{left}$, $w_{right}$ are sums of sample weights at each branch and $w_c$ is sum of weights for a particular class $c$.

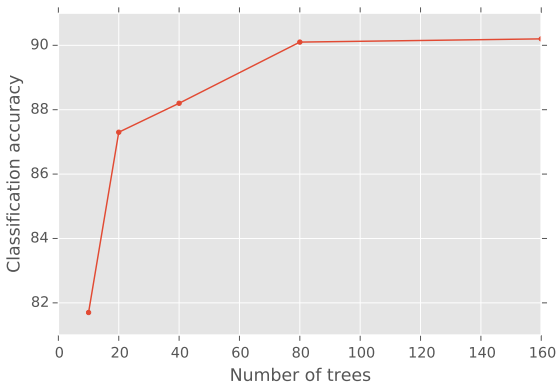
$$H(X) = -\sum_c w_c log(w_c) \tag{2}$$

Only a random subset of features (*i.e.* single pixel values at 64 different locations of the miniatures) is tested at each node of each tree and a simple threshold on this value is used to divide the data points into the left or the right partition. This helps to make the trees in the forest different from each other which leads to better generalization.

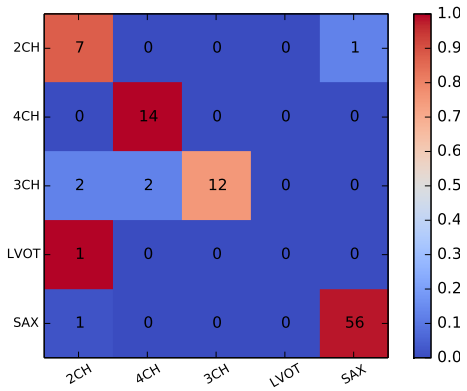
When classifying a new image, features chosen at the training stage are extracted and the image is passed through the decisions of the forest (fixed in the training phase) to reach a set of leaves. Class distributions of the reached leaves are averaged across the forest and the most probable label is selected as the image view.

## 2.4 Augmenting the dataset with jittering

Our dataset contains in total 960 image slices from 100 cardiac patients (SAX:540, 4CH: 140, 2CH: 112, 3CH: 107, LVOT: 9). This might seem rather modest compared to recent image recognition work *e.g.* [4] where deep models are trained with 15 million images. The object of interest is in general placed in the image center, however deviations exist. To account for these differences we expand the training set with extra images by applying a set of transformations of the original training images. Similarly to [4] these are translations (all shifts in x and y between -10 and 10 pixels for a 5x5 grid), but we also add rotations (angles between -10 and 10 degrees with 20 steps) and scales (1-1.4 zoom factor with 8 steps while keeping the same image size) resulting in a total of 51894 training images. Note that this extra expense is mainly at the training time. The test time is almost unaffected except that now a deeper forest can be learnt.




(a) The effect of the number of trees.




(b) Confusion matrix from one fold using image features only.

Figure 4: We selected the optimal parameters via cross validation. Despite the sample weighting the LVOT view recognition suffered from being underrepresented.

# 3 Validation

**Datasets** We validated this method on a dataset of 100 patients from multi-center study on post myocardial infarction hearts DETERMINE [3]. These consist of steady state free precession magnetic resonance acquisitions.

**Results** We ran a 25-fold cross validation by dividing the dataset on patient identifier to prevent biasing our results. This means that images from the same patient (despite being a different view) and therefore the same acquisition could never appear in the training and testing set at the same time. Here we present results for the merged short axis view labels by using features extracted exclusively either from image-based features or from DICOM slice orientation information.

| | Image features only | | DICOM orientation features only |
|---|---|---|---|
| | No jitter | With jitter | |
| Accuracy | $87.1 \pm 1.0\%$ | $90.2 \pm 1.6\%$ | $98.8 \pm 5.4\%$ |

Increasing the number of trees from 10 to 160 improves the prediction accuracy, however at the expense of increased training and classification computational cost . Beyond 80 trees the forest reaches a plateau with no further benefit (See Fig. 4(a)).

# Conclusion

In this paper we presented an automatic cardiac view recognition technique from magnetic resonance images based on image appearance and DICOM plane orientation information. The DICOM orientation based features can be used for our population with confidence when this information is present. When this is not the case, the pure image information method can be used instead. It is achieving precision on par or better with pure image based results of [10] while being much simpler and not requiring to train any landmark detectors.

By jittering the dataset with extra translations, rotations and scales we observed noticeable benefits to view recognition. Extra jittering steps might be useful to improve robustness

to intensity differences between different acquisitions. If enough examples are collected, this method can be simply extended to other pathology specific views such as in congenital heart diseases and acquisition sequences different from SSFP.

We plan to use our method to organize collections of images, semantic image retrieval and parsing of medical literature.

**Acknowledgments.** This work was partly supported by Microsoft Research through its PhD Scholarship Programme and ERC grant MedYMA. We used data and infrastructure made available through the Cardiac Atlas Project (www.cardiacatlas.org) [2][3]. This work uses scikit-learn toolkit[7].

# References

[1] Leo Breiman. Random forests-random features. pages 1–29, 1999.

[2] CG Fonseca, M Backhaus, DA Bluemke, RD Britten, JD Chung, BR Cowan, ID Dinov, JP Finn, PJ Hunter, AH Kadish, DC Lee, JAC Lima, P Medrano-Gracia, K Shivkumar, A Suinesiaputra, W Tao, and AA Young. The Cardiac Atlas Project- an Imaging Database for Computational Modeling and Statistical Atlases of the Heart. *Bioinformatics*, 27(16):2288–2295, 2011.

[3] Alan H Kadish, David Bello, J Paul Finn, Robert O Bonow, Andi Schaechter, Haris Subacius, Christine Albert, James P Daubert, Carissa G Fonseca, and Jeffrey J Goldberger. Rationale and design for the Defibrillators to Reduce Risk by Magnetic Resonance Imaging Evaluation (DETERMINE) trial. *Journal of cardiovascular electrophysiology*, 20(9):982–7, September 2009. ISSN 1540-8167.

[4] Alex Krizhevsky, I Sutskever, and GE Hinton. ImageNet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*, 2012.

[5] M Otey, Jinbo Bi, S Krishna, and Bharat Rao. Automatic view recognition for cardiac ultrasound images. In *International Workshop on Computer Vision for Intravascular and Intracardiac Imaging*, pages 187–194, 2006.

[6] J. H. Park, S. K. Zhou, C. Simopoulos, J. Otsuki, and D. Comaniciu. Automatic Cardiac View Classification of Echocardiogram. *IEEE 11th International Conference on Computer Vision*, pages 1–8, 2007.

[7] F Pedregosa, G Varoquaux, A Gramfort, V Michel, B Thirion, O Grisel, M Blondel, P Prettenhofer, R Weiss, V Dubourg, J Vanderplas, A Passos, D Cournapeau, M Brucher, M Perrot, and E Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[8] A M Taylor and J Bogaert. Cardiovascular MR Imaging Planes and Segmentation. In Jan Bogaert, Steven Dymarkowski, Andrew M Taylor, and Vivek Muthurangu, editors, *Clinical Cardiac MRI SE - 333*, Medical Radiology, pages 93–107. Springer Berlin Heidelberg, 2012. ISBN 978-3-642-23034-9.

[9] Antonio Torralba, Rob Fergus, and William T Freeman. 80 Million Tiny Images: a Large Data Set for Nonparametric Object and Scene Recognition. *IEEE transactions on pattern analysis and machine intelligence*, 30(11):1958–70, November 2008. ISSN 1939-3539.

[10] Y Zhou, Zhigang Peng, and XS Zhou. Automatic view classification for cardiac MRI. In *9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1771–1774, Barcelona, 2012. IEEE. ISBN 9781457718588.

# Automatic Identification of Early Miscarriage Based on Multiple Features Extracted From Ultrasound Images

Shan Khazendar[1]
Shan.khazendar@buckinghamac.uk
Jessica Farren[2]
Jessicafarren@gmail.acom
Hisham Al-Assam[1]
Hisham.al-assam@buckingham.ac.uk
Hongbo Du[1]
Hongbo.du@buckingham.ac.uk
Ahmed Sayasneh[2]
Sayasneh@imperial.ac.uk
Tom Bourne[2]
Womensultrasound@btinternet.com
Sabah A.Jassim[1]
Sabah.jassim@buckingham.ac.uk

[1] The University of Buckingham, Applied Computing Department, Buckingham, MK18 1EG ,UK.

[2] Queen Charlotte's and Chelsea Hospital, Department of Early Pregnancy, Imperial College London, W12 0HS, UK.

**Abstract**

Ultrasound is one of the most widely used multipurpose imaging modalities that is ideal for monitoring and diagnosing early pregnancy events. The first sign and measurable element of an early pregnancy is the Gestational Sac (GS). Currently, the size of GS is manually measured from an ultrasound image of the GS. This paper argues that the Mean Sac Diameter (MSD) derived from the manual measurements results in inter- and intra-observer variations, which may lead to difficulties in diagnosis. The paper proposes a fully automated diagnosis solution to accurately identify miscarriage cases in the first trimester of pregnancy based on currently used MSD as well as alternative geometric features extracted from the image. Our experimental results show that the perimeter and volume of the GS are effective features where the perimeter can outperform the MSD. Furthermore, our study shows that the identification accuracy of early miscarriage can be further improved by combining the perimeter, volume and MSD of the GS.

## 1 Introduction

Medical imaging techniques have been increasingly deployed in the past decades to diagnose various types of diseases. Among medical imaging modalities, ultrasound imaging is considered to be safe, non-invasive, portable, accurate, and cost effective. These advantages have made the ultrasound imaging the most common diagnosis tool deployed hospitals around the world [1]. Ultrasound imaging is also considered as an effective modality particularly for monitoring pregnancy because of its safety without hazard of radiation [2].

The first three months, known as the first trimester, are the most crucial period in pregnancy. Monitoring pregnancy within this period enables clinics to evaluate the development, growth, and wellbeing of the foetus [3]. The first sign and measurable

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

element of an early pregnancy is the Gestational Sac (GS). The American College of Radiology guideline defines miscarriage as being an empty GS with a Mean Sac Diameter (MSD) greater than or equal to 16 mm [4, 5]. However, the Royal College of Obstetricians and Gynaecologists of the United Kingdom first recognised miscarriage as being an empty sac with MSD greater than or equal to 20 mm, but later concluded that an empty GS with MSD greater than or equal to 25mm should be introduced as a new guideline to minimize the risk of false positive diagnosis of miscarriage [6].

Estimating the size of the GS is currently done manually. The manual process involves multiple subjective decisions when three diameter measurements on the GS are taken from an ultrasound image to establish the MSD [6]. The subjective decisions increase the inter- and intra-observer variations which may lead to difficulties and even errors in the diagnosis stage. Therefore an automated way of measuring the size of the GS from a given ultrasound image is desirable. Unlike other types of medical image, ultrasound images are corrupted by speckle noise that tends to reduce image resolution and contrast, and consequently reduce the diagnostic value of the image. Therefore speckle noise reduction is an important requirement whenever ultrasound imaging is used.

A large amount of research into ultrasound image de-nosing has been undertaken [7,8,9], but the research in the area of gestational sac segmentation and enhancement is very limited. In [10], Chakkarwar et al. presented an automatic method for GS segmentation using a database of 12 images with average accuracy of 83.3%. Their method starts by de-speckling the image using a combination of contrast enhancement, low pass filter and wiener filter, followed by thresholding. Zhang et al. [11] proposed a three-step automatic detection method for GS from a video using AdaBoost method to detect the GS in each frame followed by an efficient method exploiting the local context to reduce false positive detection. The algorithm was tested on 31 videos and achieved a GS detection rate of 87.5%.

This paper proposes an automatic solution for accurate segmentation and quantification of GS. Besides MSD, the proposed solution also extracts some geometrical features from GS images such as volume, perimeter, area, circularity, compactness, solidity and eccentricity from two most important planes - transverse and Sagittal for GS that may be relevant to miscarriage identification. We have used k Nearest Neighbor (kNN) classifier to identify early miscarriage cases based on the extracted features. To verify the performance of the proposed solution, we have conducted experiments on a data set of 68 ultrasound images. Our experimental results show that the proposed solution can achieve a higher level of accuracy using the perimeter compared with the MSD. We also showed that combining the perimeter, volume and MSD features can further improve the accuracy of miscarriage diagnosis.

# 2 Materials and Methods

## 2.1 Materials

A database of 68 ultrasound images for GS was prepared to cover both Normal (44) and Miscarriage (24) cases. Each image was captured from two planes - transverse and sagittal for GS. The database was labeled by experts in early pregnancy department, Queen Charlotte's and Chelsea Hospital, London, UK.

## 2.2 Methods

Figure 1 shows the block diagram of the underlying process of the proposed solution for automatic identification. Each stage of the process will be explained in detail in the following sub-sections.

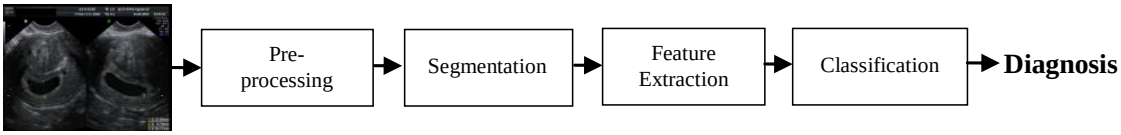



Figure 1: Block Diagram of the Major Steps of the Proposed Method

### 2.2.1 Pre-processing

Ultrasound images of GS are very dark and hence the accuracy of GS segmentation can be badly affected. We first separate the sectional images as shown in Figure 2 (a) and (b), and then brighten each image by using the following adaptive method: if the mean of all pixels intensity values of the image >=55, the intensity value of each pixel is multiplied by 2; otherwise it is multiplied by 4. Unlike histogram equalisation, this transformation gives more weight for the dark pixels (where the detailed information lays) by stretching them over the whole grey-sale range. The enhanced image is shown in Figure 2 (c).

### 2.2.2 GS Segmentation

Step 1: A simple method to extract an object from its background is through thresholding. The Otsu method [12] is typically used to automatically select the best threshold. However, this method did not perform well on most of the images in our database because our database contain images where are very different intensity. Therefore, we chose the mean pixel values of the original image as an alternative threshold. This simple solution is effective on almost all type of images (bright and very dark) in our database as illustrated by the example in Figure2 (d).
Step 2: A median filter with a window size of 15x15 is applied to smooth the boundary of the binary object without losing its original shape in addition to filling holes and removing small outlier regions. The result is shown in Figure2 (e).
Step 3: The best-fit ellipse is then found to estimate the shape of the GS due to the fact that GSs are usually round in the early pregnancy, but as the sac grows it often appears more elliptical. Figure 2(f) shows the ellipsoid shape on the GS
Step 4: To extract the border of GS, the morphological operation erosion is first applied. The resulting image is then subtracted from the binary image from step 2. The border of the GS is illustrated in Figure2 (g).
Step 5: The border extracted in step 4 is superimposed on the original image to clearly visualize the segmented GS as shown in Figure2 (h).

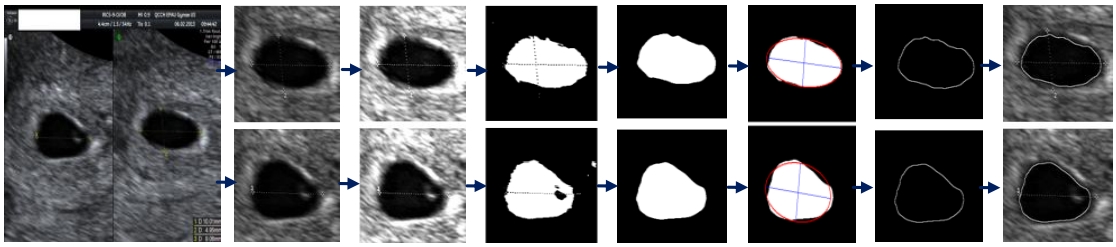
### 2.2.3 Feature Extraction

As explained earlier, each GS is visualised in two perpendicular sections/planes. Our algorithm first finds the best fitting ellipse for the segmented GS in each plane. Assuming the GS has ellipsoid shape in 3D, the three principal axes of the ellipsoid can be estimated by the major axis (A), minor axis (B) of the ellipse from the first plane and the depth (C) from the second plane. After that, we extracted the following geometric features from each GS:

1. The Mean Sac Diameter (MSD) : Based on the three principal axes A, B, and C, the MSD is given by:
$$MSD = \frac{A+B+C}{3}$$
2. Volume: The volume of the GS can be estimated using the three semi-principal axes as follows:
$$Volume = \frac{4}{3}\pi\ ABC$$
3. Perimeter: It can be accurately calculated by simply counting the number of pixels around the boundary of the GS, and then taking the average of the perimeter from both sections to produce a single perimeter measure.

### 2.2.4 Classification

At this stage, a classifier is normally derived from a set of examples. In principle, any suitable classifier can be trained and deployed in this step. In this study, we particularly used a simple kNN classifier to test the effectiveness of the extracted features.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 2: (a) Original Image in two sectional, (b) Cropped image GS (c) Enhanced Images, (d) Binarized image (e) Border of the GS smoothed (f) Border extracted (g) Superimposed the border on the original image (h) Result of segmented GS with best fitted ellipsoid

# 3 Experiment Results & Discussion

## 3.1 Experiments and Results

In our experiments, a stratified cross-validation was employed and a random sample was drawn 15 times. For each sample, 24 random normal and 24 miscarriage cases were chosen. We report the average accuracy, the average sensitivity and the average specificity. The results in Figure 3 illustrate that when the automatically extracted MSD feature is used, we obtained an overall accuracy of 98% with sensitivity (miscarriage) 97% and specificity (normal) 99%. When the Volume feature is used, the level of accuracy is marginally lower with a larger discrepancy between sensitivity 95.5% and specificity 100%. The Perimeter feature has achieved a better sensitivity of 98.6% while specificity remains at 99%. The difference between the sensitivity and specificity has also been narrowed, indicating an overall improvement over the MSD and the Volume features. Perimeter is a more robust feature which is automatically measureable from each single image of the GS. Furthermore, the results in figure 3 also show that combining the three features will further improve the performance (accuracy, sensitivity and specificity) to 99.5% through complementing the strengths of the individual features.

## 3.2 Discussion

In order to further understand the applicability of the extracted features in classification, we also tested the features using two other types of classifiers: decision tree (DT) and

support vector machine (SVM) as well as kNN [13] when k=1. Figure 4 shows a consistent pattern of performance of MSD and perimeter across the classifiers, but the performance of volume feature deteriorates for SVM classifier and for kNN , k=1 classifier, due to some border line cases which make the decision very sensitive.

We also have investigated into other features such as area, compactness, circularity, solidity and eccentricity [14]. Area is another indicator of GS size and is calculated by first finding the actual number of pixels in the GS region (GS) and then converting the number of pixels to area measurement in $mm^2$. The classification test on this feature shows an inferior performance compared to the other three features. Compactness, solidity, circularity and eccentricity are indicators of roughness, irregularity, and the shape of GS borders. Our experimental result shows that these features have little diagnostic value for miscarriage cases except the area. Figure 5 illustrates the classification results for these features. The reason for the poor performance of these features may well be due to the images in our data set where most sacs, miscarriage or normal, have regular and smooth borders. Initial investigation on replacing the MSD with the three measurements , major axis (A), minor axis (B) of the ellipse from the first plane and the depth (C) from the second plane separately to be fed into the classification scheme did not show any improvement over the MSD. However, more images of various border characteristics are needed in future investigations to draw solid conclusions.

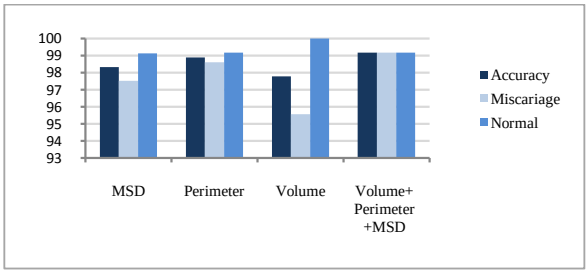



Figure 3: Comparing miscarriage classification accuracy, sensitivity and specificity based on MSD, perimeter, volume and multi features using kNN, k=3

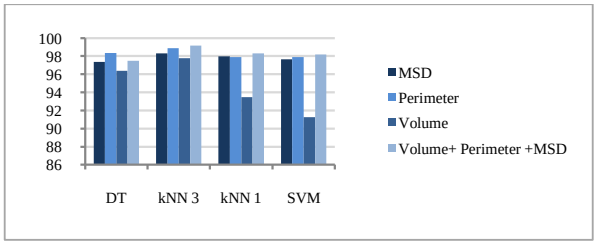



Figure 4: Classification accuracy based on MSD, perimeter, and volume using 4 different Classifiers

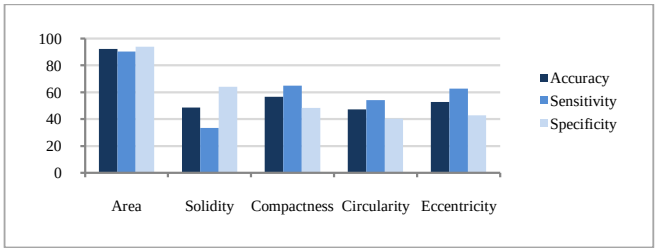



Figure 5: The system performance based on area, compactness, circularity, solidity and eccentricity using kNN, k=3

# 4 Conclusions

In this paper, an automatic computer based solution for segmenting, quantifying and classifying the GS has been proposed for miscarriage diagnosis in a very early stage of pregnancy. Firstly, our results show that our preprocessing task was successful for segmenting the GS. Secondly, we have investigated the effectiveness of new features such as perimeter and volume to further improve the classification accuracy. The result shows that using perimeter as a single feature outperforms the use of MSD in terms of miscarriage classification accuracy. Thirdly, our study shows that the combining perimeter, volume and MSD can improve the classification accuracy even further. Finally, we establish that the contribution of other features such as area, solidity, circularity, compactness, and eccentricity in diagnosis is very limited. Our future work includes segmentation, quantification and classification of other types of miscarriage cases such as GS with Yolk or Embryo inside.

# 5 References

[1] O. Michailovich and A. Tannenbaum, Despeckling of medical ultrasound images. *IEEE Trans Ultrason Ferroelectr Freq Control*, 53(1): p. 64-78, 2006.

[2] R. Geirsson and R.Busby-Earle, Certain dates may not provide a reliable estimate of gestational age. *Br J Obstet Gynaecol*,. 98(1): p. 108-9, 1991.

[3] A. Kaur,Transvaginal ultrasonography in first trimester of pregnancy and its comparison with transabdominal ultrasonography. *J Pharm Bioallied Sci*, 2011. 3(3): p. 329-38.

[4] T. Bourne and C. Bottomley, When is a pregnancy nonviable and what criteria should be used to define miscarriage?*Fertil Steril*, 98(5): p. 1091-6, 2012.

[5] C. Levi, E. Lyons, and D. Lindsay, Ultrasound in the first trimester of pregnancy. *Radiol Clin North Am*, 28(1): p. 19-38,1990.

[6] Ectopic pregnancy and miscarriage, Diagnosis and initial management in early pregnancy of ectopic pregnancy and miscarriage. *NICE clinical guideline* 154, 2012.

[7] J. Awad, Prostate segmentation and regions of interest detection in transrectal ultrasound images. *Diss. University of Waterloo*, 2007

[8] P. Hiremath, J. Tegnoor, and S. Badiger. Despeckling Medical Ultrasound Images using the Contourlet Transform. *IICAI*. 2009.

[9] P. Hiremath and J. Tegnoor, Automatic detection of follicles in ultrasound images of ovaries using edge based method. *IJCA, Special Issue on RTIPPR*,2: p. 120-125, 2010.

[10]A. Chakkarwar,S. Joshi, and S. Revankar. Automated analysis of gestational sac in medical image processing. in *Advance Computing Conference (IACC), IEEE ,2010*.

[11] L. Zhang,S.Chen,S.Li and T.Wang. Automatic measurement of early gestational sac diameters from one scan session. *in SPIE Medical Imaging. International Society for Optics and Photonics*,2010.

[12] N. Otsu, A threshold selection method from gray-level histograms. *Automatica*, 11(285-296): p. 23-27, 1975.

[13] H. Du, Data Mining Techniques and Applications: An Introduction, Course Technology Cengage Learning, 2010.

[14] F. Olver, NIST handbook of mathematical functions: Cambridge University Press,2010.

# Challenge Abstract

# Challenges in Functional Imaging of the Visual Cortex: Data-driven Model-agnostic Estimation of Population Receptive Fields

Francesca Pizzorni-Ferrarese

Jonas Larsson

Department of Psychology
Royal Holloway, University of London
Egham, UK

## Abstract

Disorders that affect sight and the visual system can afflict people across the life span. Common such disorders are amblyopia (a disorder in young people characterised by degradation in the visual input from one eye to the visual cortex during critical periods of development, leading to deficits in stereopsis and spatial vision) and age-related macular degeneration (a disorder characterised by a progressive loss of vision at the centre of gaze). A challenge for the development of better treatments for these disorders is that the underlying physiological mechanisms are not completely understood. A key problem that has limited the study of these mechanisms is that existing methods for investigating human visual cortical function - primarily functional magnetic resonance imaging (fMRI) - cannot easily be used with patients, largely as a result of symptoms associated with the disorders. A main limitation is represented by the requirement for steady fixation and relatively long data acquisition times. Because patients with visual disorders, often have severe fixational instability, these methods cannot be straightforwardly applied to such patients. Moreover, deficits in attention and motor control, as well as issues related to the age of patients, make long scanning sessions impractical. Hence there is a need for development of methods that are robust to fixation instability and subject motion, while being practically feasible in a relatively short time span.

To this end, population receptive field (pRF) modelling ([1], [2], [5] and [3]) is a relatively novel approach that seeks to characterise individual fMRI voxels both in terms of their spatial response properties and visual field representation, as well as their stimulus selectivity. pRF modelling, while still measuring responses reflecting the pooled activity of thousands of neurons, is conceptually more similar to neuronal recording methods used to study animal models of visual system disorders, and thus holds promise for providing a better understanding of the neuronal mechanisms of the disorders by linking measurements in humans and animals. A limitation of current pRF modelling techniques is their requirement of assumptions about the receptive field structure that puts some a priori constraints on the ability to extract the pRF. To overcome these problems, we propose a new data-driven model-agnostic[1] approach that estimates the structure of the pRF. Our method does not make a priori

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

[1] It forsakes the need for a model and requires no a priori knowledge, as the inferences are informed by the data through the layers of information that are available in the intrinsic salient structure of the dynamic response.

assumptions about the specific haemodynamic response function (HRF) and pRF shapes and is therefore a useful tool for uncovering the underlying pRF structure at different spatial locations in an unbiased way. Our method is based on the idea of reverse correlation ([4]) applied to BOLD signals. Reverse correlation or white noise analysis is a standard technique for estimating receptive fields of single neurons by cross-correlating the stimulus with the spiking response for a neuron, using so-called spike triggered averaging. A limitation of applying this technique to estimate pRFs from fMRI data is that it ordinarily requires large numbers of spatially independent stimuli to avoid biased estimates. In this work we have applied the reverse correlation approach by making specific assumptions about the temporal structure of the fMRI response. Figure 1 outlines the whole processing workflow, from the raw data to the final estimation of the pRF and time course prediction

Our method provides (i) rapid and reliable estimates of the size and shape of pRFs with fewer stimuli than conventional reverse correlation, (ii) does not have as stringent requirements for spatially independent stimuli, and (iii) has the potential to be used in patients with visual disorders that cannot be routinely tested with currently available methods.

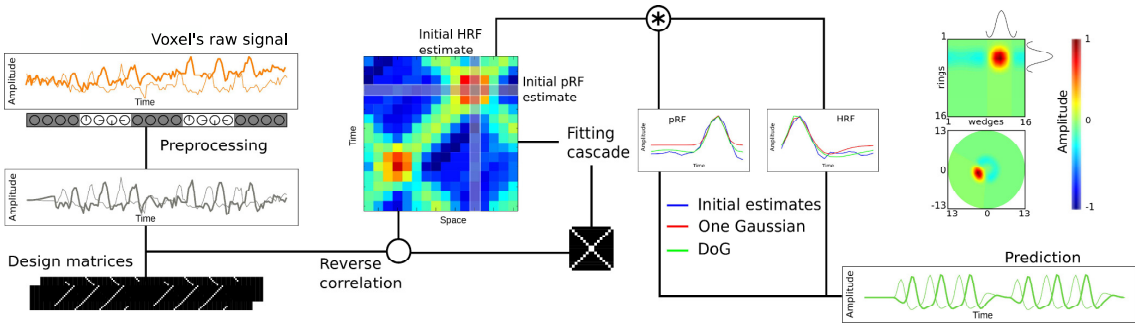



Figure 1: **pRF estimation workflow.** The reverse correlation is applied to the filtered fMRI time courses. The spatio-temporal kernel obtained provides initial estimates for the pRF and HRF. These are subsequently refined through a fitting cascade process.

# References

[1] Serge O Dumoulin and Brian A Wandell. Population receptive field estimates in human visual cortex. *Neuroimage*, 39(2):647–660, 2008.

[2] J Larsson and D J Heeger. Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci*, 26(51):13128–13142, 2006.

[3] S Lee, A Papanikolaou, N K Logothetis, et al. A new method for estimating population receptive field topography in visual cortex. *NeuroImage*, 81:144–157, 2013.

[4] D Ringach and R Shapley. Reverse correlation in neurophysiology. *Cognitive Sci*, 28 (2):147–166, 2004.

[5] W Zuiderbaan, B M Harvey, and S O Dumoulin. Modeling center-surround configurations in population receptive fields using fMRI. *J Vision*, 12(3):10, 2012.

# Poster Session 1: Segmentation and Shape Analysis

# Automatic Ribs Segmentation and Counting from Mouse X-ray Images

Omar Al Okashi[1]
Hongbo Du[1]
Joanne Selway[2]
Chris Lelliott[3]
Simon Maguire[3]
Dave Melvin[3]
Sanger Mouse Genetics Project[3]
Kenny Langlands[2]
Hisham Al-Assam*[1]
(hisham.al-assam@buckingham.ac.uk)

[1] Department of Applied Computing, University of Buckingham, MK18 1EG
[2] Buckingham Institute for Translational Medicine, University of Buckingham, MK18 1EG
[3] Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA

**Abstract**

The automatic detection and quantification of skeletal structures has a variety of applications for biological research. This paper proposes an automatic solution for rib segmentation and counting based on structural properties of ribs in mouse X-ray images. The solution consists of five stages, including alignment, cropping the region of interest, image enhancement, rib segmentation, filtering of non-rib objects and rib counting. Experimental results on a set of 100 X-ray images showed consistent good performance of the algorithm compared with a manual annotation by domain experts, with an overall accuracy of about 88%.

# 1 Introduction

The Mouse Genetics Project [1] was initiated to systematically knock-out mammalian genes and screen for a broad range of resulting traits. One element of the phenotyping pipeline is the thorough evaluation of the skeleton through systematic X-ray imaging of individual animals and manual observations of variation. Given the volume of images that are generated by high-throughput screening, an automated approach to annotation and triaging of images will reduce time and human resource requirements and possibilities of human error.

Accurate identification of any perturbation in skeletal structure is of significant biological relevance. For example, human congenital abnormalities can lead to additional ribs and potential nerve damage to the arms (i.e. thoracic outlet syndrome), whereas loss of ribs can lead to damage within the thoracic cavity due to a lack of protection [2]. Therefore, the identification of genes associated with such abnormalities will provide insight into both specific and generalised abnormalities in skeletal structure.

Our aim is to exploit image processing techniques to enable the automatic counting of ribs from X-ray images of mice. However, this task faces the following specific

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

challenges. First, the source images in our study are full-body images and hence segmentation of regions of interest (ROIs), i.e. the rib areas, is essential. Second, ribs can be difficult to distinguish in a 2D X-ray image because of their low, sometimes extremely low, colour intensity. Third, most ribs are connected to each other on the edge of the ribcage, making automatic separation a difficult task. Furthermore, the inherent noise in X-ray images presents technical difficulties in rib identification.

Previous work has described rib segmentation and positioning in humans for different medical modalities. The method presented in [3] exploits the capability of Hough Transform to detect lines that determine the first rib in the sagittal plane. The position of the first rib was then used as a reference to determine the initial position of other ribs. Li et al. [4] presented a method that firstly determined the position of the spinal canal after thresholding the image (empirically determined), and then the location for a seed region for every rib. These seeds were then used as a starting point to segment ribs from vertebrae. A model based segmentation was used in [5] to segment and label the ribcage. Triangulated surface models of all 24 ribs were initially created.After that, an adaptation of these models to their corresponding anatomical object to form an initial ribcage was done by using shape constrained model. To determine a proper position for a complete model, a ray based search approach was used to detect a ribs. However, the methods described in [3], [4] and [5] were applied to segment ribs in 3D CT images, and it remains unclear if these methods were applicable to single view X-ray images.

Rib segmentation in X-ray images was considered in [6]. The proposed solution first smooths the image using a Gaussian function, followed by passivation masking, i.e. subtracting the original from the smoothed image. The wavelet image fusion between original and segmented image has been used to enhance the image. After that the Otsu segmentation method [7] has been used to detect lung parenchyma. Finally, ribs were denoised using wavelet and morphology operations such as corrosion and expansion. However, the final result shows connected ribs as well as noises from chest tissues and other organs which makes it unsuitable for automatic counting of the ribs.

This paper proposes a fully-automated method to segment and count ribs in X-ray images. Experimental results on a data set of 100 manually annotated X-ray images confirmed the effectiveness of the proposed method with an overall accuracy of 88%.

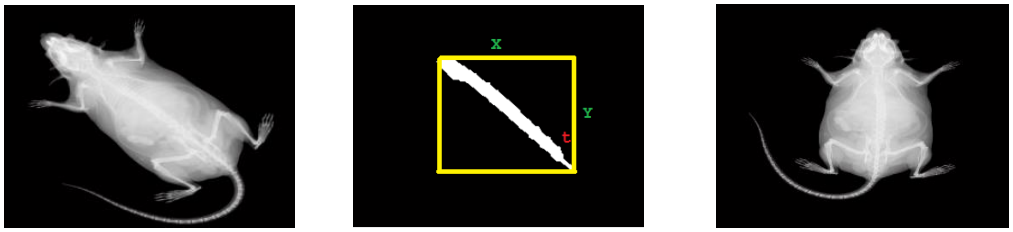
# 2 Proposed Method

Our proposed solution consists of a sequence of 5 steps. The process started with the automatic alignment of a dorsoventral 2D X-ray image to orientate the mouse skeleton in a vertically-upright position. This stage is essential for correct detection of rib ROIs, and to allow later exploitation of left and right ribcage symmetry. After that, a mask image of reference points is used to crop out an ROI containing ribs. This is then followed by image enhancement in spatial and frequency domains to improve the contrast of the cropped area. The next step is the segmentation of ribs from other organs in the chest. The last step is removal of unwanted objects and counting of ribs.

## 2.1 Automatic Alignment

To perform alignment of the mouse skeleton as shown in Figure 1-a, the spine must be correctly located. We realised that the spine has higher contrast than other surrounding tissues and bone structures. To achieve the objective of spine location, we first applied

adaptive histogram equalization (ADHE) [8] to enhance the contrast, followed by Gamma transformation [9] to highlight the high contrast bones and eliminate other low intensity structures . Morphological operations are then applied to make the spine a complete object without breaks, and to filter out unwanted objects. Once the approximate spine was located, the rotation angle required for the vertical alignment is determined from the bounding box dimensions as illustrated in Figure 1-b using the inverse tangent function. The result of the rotation is shown in Figure 1-c.




(a) (b) (c)

Figure 1: Alignment process:(a)Input image;(b) Detected spine;(c) Result of alignment.

## 2.2 Cropping of ROI

To crop the ROI, we need to locate relevant x and y coordinates of the region. It is realised from our observation of the images that such coordinates can be estimated from the x and y coordinates of the limbs. To obtain an image of limbs, we first apply a High Emphasis-Butterworth High Pass Filter (HE-BHPF) [9]. HE-BHPF uses a multiplier to emphasise the high frequency components such as higher contrast bone, limbs and skull while maintaining the low frequency for other organs, as shown in Figure 2-a. This will enable a segmentation of these bones in the next step. After that, the regiongrow segmentation technique that segments different objects by grouping pixel with similar properties (e.g. Intensity) in a region [9] is used to extract high contrast objects such as limbs and skull (shown in Figure 2-b). Once the limbs are segmented, we locate the x and y coordinates of the limbs as shown in figure 2-b, and then use the coordinates to crop the ROI from the original image as shown in figure 2-c.

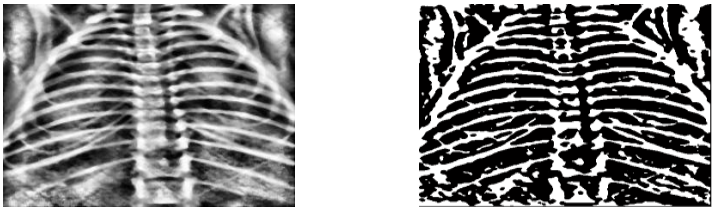
(a) (b) (c)

Figure 2: Cropping of ROI: (a) Original image processed with the HE-High-Pass filter; (b) Image of processed with region growing segmentation; (c) Cropped ROI.

## 2.3 ROI Image Enhancement

The cropped ROI may be of very low contrast that may lead to some ribs, particularly the lower ones, being missed by the rib segmentation step. To increase rib contrast, we apply image enhancements through the following two steps:

1. Enhancement in the spatial domain. Similar to the spine location, ADHE is used to enhance the contrast of the overall image first. A median filter is then applied to suppress background noise and smooth the image as shown in Figure 3-a.
2. Enhancement in the frequency domain. It can be seen from Figure 3-a that some overlapping between the ribs and the lung tissues is still present despite the enhancement in spatial domain. For this reason, we decide to further process the enhanced image in the frequency domain by applying first the Butterworth High-Pass filter (BHPF) to highlight and enhances image details, and then a median filter to suppress noise. Figure 3-b shows the result of this step.



(a) (b)

Figure 3: Enhaning the cropped image: (a) after adaptive HE and median filters, and (b) after applying BHPF and median filters.

## 2.4 Rib Segmentation

To segment ribs from other structures, we need to remove the spine and the edge of the ribcage from the enhanced ROI. We use the spine outline shown in Figure 1-b as a mask to clear the spine, and split the ROI into two parts as shown in Figure 4-a.

To remove the edge of the ribcage, we proposed a linear function that initiates from the edge of the top rib towards a lower one. This function trims most of the noisy edges of the ribcage and separate the ribs (Figure 4-b). Although crude in cutting off parts of ribs, as we are only concerned with rib counting as opposed to rib length measurement, the function is effective for separating the edge of the ribcage from disjointed ribs.

(a) (b)

Figure 4: Rib Segmentation process: (a) Spine removal ;( b) Ribcage removal and splitting

## 2.5 Filtering Out Non-Rib Objects and Counting Ribs

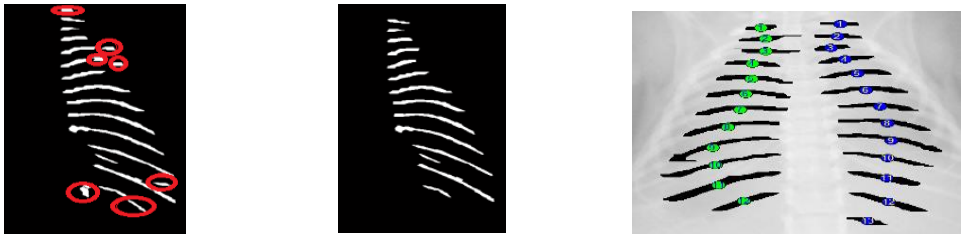
It can be seen that the image in Figure 4-b can still not be used directly for rib counting as the image has many non-rib objects that can affect the counting accuracy. By analysing the properties of the ribs, we use the following criteria to remove non-rib objects:

1- Size. If the size of any object is smaller than a specific threshold (i.e. the size of the smallest rib), it will be marked as noise.
2- Shape. Ribs have well-defined shape properties. We approximate each object as shown in Figure 4-b by the smallest ellipse that contains the object, and measure the ratio of

the minor to major axis. Since any rib should have its major axis value greater than its minor axis value, we can eliminate circular, vertical and irregularly-shaped objects.

3- Orientation. After the initial alignment, ribs have expected orientation angles, e.g. a rib cannot be vertical. Therefore, setting a rational threshold for the orientation of left and right ribs can eliminate many non-rib objects.
4- Starting position. All ribs start approximately at the same spatial x-coordinate. Using the mean x-coordinate for the start position of all left and right groups of ribs, non-ribs objects that have start position far away from the mean value, e.g. $2\sigma$, can be removed.

Figure 5 shows the effectiveness of applying these criteria on an example image. Non-rib objects in the circles in Figure 5-a are removed. After filtering, the image is ready for rib counting. This is done by simply scanning along the y-coordinates as shown in Figure 5-c.



(a) (b) (c)

Figure 5: Filtering-out non rib objects: (a) Before filtering, with non-ribs circled; (b) Filtering using rib properties; (c) Final annotated image.

# 3 Experiments and Discussion

We have tested our method on a dataset of 100 full body dorsoventral view X-ray images that were previously manually annotated by domain experts. These 100 images include a variety of images of different degrees of difficulties in image analysis. Since the proposed solution goes through three main stages, i.e. body alignment, ROI cropping, and rib segmentation and counting, we evaluate the accuracy at each of the three stages separately. Automatic body alignment is considered accurate if the position of the body is in a vertical position by observation.The results show that our automatic alignment achieves an accuracy of 98%. The ROI cropping is considered accurate if the cropped area isolates all the ribs without missing any. Our method cropped 98 out of 98 correctly aligned images, i.e. 100% accuracy. The two inaccurately aligned images has caused error at the cropping stage where the top rib is not included in the area as shown in Figure 6-a.

(a) (b)

Figure 6: Examples of inaccurate counting: (a) Missed upper (b) Merged ribs.

Out of the 98 images that were correctly cropped, ribs in 86 of them (i.e. 87.8%) were counted accurately according to the manual annotations. The other 12.2% of images

contain missing, merged or discontinuous ribs. Missing ribs occur due to extremely low intensity of lower ribs, or errors in removing ribcage edges. Merged ribs, as shown in Figure 6-b, can be a result of reflection due to the noisy nature of X-ray images.

# 4 Conclusions

Segmentation and counting of ribs is essential for assessing skeletal phenotype. We proposed a fully-automated algorithm that segments and counts ribs in X-ray images. We have demonstrated a comparable accuracy with existing manual methods, and shown the ability of our solution to overcome a variety of challenges presented by the segmentation of X-ray images. The aim of this study was to locate and count ribs rather than to segment and identify them in their entirety.We are now working to improve the level of accuracy of the alignment, and the removal of non-rib objects by exploiting more sophisticated filters and investigating various thresholds used in this process. Our future work also includes a more in-depth analysis for the proposed schemes on a larger set of images.

# 5 Acknoweledgment

We thank the Sanger Research Support Facility, ES Cell Mutagenesis, Transgenic Technologies and Mouse Informatics Groups for support in providing the mice, X-Rays and images used in this paper.

# 6 References

[1] J. White,G. Anna-Karin, N. Karp, E. Ryder, M. Buljan, J. Bussell, J. Salisbury, S. Clare, NJ. Ingham , C.Podrini, R .Houghton , J.Estabel, JR.Bottomley, D.Melvin,D.Sunter ,N.Adams; D.Tannahill, D .Logan, D .Macarthur, J.Flint , V.Mahajan, SH.Tsang ,I. Smyth, F.Watt , W.Skarnes , G.Dougan , D.Adams , R.Ramirez-Solis ,A.Bradley , KP .Steel . Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. Cell 154, no. 2, pp. 452-464, 2013.

[2] D. Wattanasirichaigoon, C. Prasad, G. Schneider , J. A. Evans and B. R. Korf, Rib defects in patterns of multiple malformations: a retrospective review and phenotypic analysis of 47 cases, *American Journal of Medical Genetics Part A*, vol. 122, no. 1, pp. 63-69, 2003.

[3] H.Li, J.Li, S. Guo, J. Liu, and Y. Kang. Automatic rib positioning method in CT images. *In Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference* , pp. 1-4. IEEE, 2010.

[4] J. Lee and A. Reeves. Segmentation of individual ribs from low-dose chest CT. *In SPIE Medical Imaging*, pp. 76243J-76243J. International Society for Optics and Photonics, 2010.

[5] T. Klinder, C. Lorenz, J.Berg, S. Dries, T. Bülow, and J. Ostermann. Automated model-based rib cage segmentation and labeling in CT images. *In Medical Image Computing and Computer-Assisted Intervention*–MICCAI 2007, pp. 195-202. Springer Berlin Heidelberg, 2007.

[6] W. Cui; T.Yu; L.Ren; H.Shao, Ribs segmentation based on image fusion and wavelet de-noising, *Biomedical Engineering and Informatics (BMEI), 2012 5th International Conference on* , vol., no., pp.362,366, 16-18 Oct. 2012.

[7] N.Otsu, A Threshold Selection Method from Gray-Level Histograms, *IEEE Transactions on Systems*, Man, and Cybernetics, Vol. 9, No.1, pp. 62-66, 1979.

[8] W. Pratt, Digital Image Processing, 3rd ed., Wiley, 2001.

[9] R. Gonzalez and R.Woods, Digital Image Processing, 3rd ed., Pearson International Edition, 2008.

# Active Contours with Weighted External Forces for Medical Image Segmentation

Alaa Khadidos[1,2]
A.Khadidos@warwick.ac.uk
Victor Sanchez[1]
vsanchez@dcs.warwick.ac.uk
Chang-Tsun Li[1]
C-T.Li@warwick.ac.uk

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

[2] King Abdulaziz University
Faculty of Computing and Information Technology, Jeddah, Saudi Arabia

**Abstract**

Parametric active contours have been widely used for image segmentation. However, high noise levels and weak edges are the most acute issues that hinder their performance, particularly in medical images. In order to overcome these issues, we propose an external force that weights the gradient vector flow (GVF) field and balloon forces according to local image features. We also propose a mechanism to automatically terminate the contour's deformation. Evaluation results on real MRI and CT slices show that the proposed approach attains higher segmentation accuracy than snakes using traditional external forces, while allowing initialization using a limited number of selected points.

## 1 Introduction

Parametric active contours, or snakes, model the movement of a dynamic curve towards an object's boundary under the influence of internal and external forces [5, 8]. Parametric snakes produce closed and smooth object boundaries while allowing the separation of objects in contact by prohibiting contour fusions. The latter is an important advantage over non-parametric approaches, such as the level-set approach [15].

Despite the outstanding results of parametric snakes for image segmentation, their accuracy is relatively low when the amount of image clutter and noise is high [14]. In such cases, the use of various external forces that help the snake to conform to the desired boundary is an appropriate solution. For example, Cohen [2] propose *balloon* forces to guide the snake to the object's boundary in a similar way a balloon inflates or deflates. Xu *et al*. [11] propose the gradient vector flow (GVF) field as an external force to help the snake to conform to concave boundaries. Other proposals include using vector field convolution (VFC) as an external force, which is calculated by convolving the image edge map with a user-defined vector field kernel [6]; integrating the GVF field with prior directional information manually provided [14]; convolving the image with a sigmoid function before computing the GVF field [12]; adaptively computing the GVF field according to the characteristics of an image region [10];and computing a mean shift based GVF (MSGVF) to balance internal and external forces and drive the curve towards the desired boundary [13].

Although balloon forces may help guide the snake to the desired boundary over noise areas, the strength of the force must be appropriately set in order for the snake to detect weak edges and avoid snake leakages. GVF-based methods, and many of its variants, are proposed in order to address shortcomings of the balloon force; these external forces, however, may still require that the initial snake be initialized close to the desired boundary to obtain accurate segmentation results, particularly in noise images [1]. This inevitably involves manually

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

selecting several initial points or snake elements. Moreover, regardless of the external force used, the appropriate number of iterations to accurately conform the desired boundary must be carefully selected in an empirical manner.

In this work, we propose an external force that combines the advantages of GVF and balloon forces, especially for cases when the snake is initialized far from the desired boundary in a noisy area and with a very limited number of snake elements. Specifically, balloon forces are used to guide the snake to the object's boundary even in the presence of image clutter and noise; while GVF forces are used to improve segmentation accuracy even around weak edges. We control the effect of these two forces by a weighting function based on local image features. We also propose a mechanism to terminate the snake's deformation process without having to manually indicate the number of iterations. The proposed approach minimizes snake leakages and considerably reduces the number of initial snake elements, making it suitable for medical image segmentation with little manual intervention.

## 2 Proposed External Force

A snake is a curve $C(s) = [x(s), y(s)]$, with $s \in [0,1]$ at position $(x,y)$, that evolves to an object's boundary by minimizing the energy functional $E_S = \int_0^1 [E_{\text{int}}(C(s)) + E_{ext}(C(s))]ds$, where $E_{\text{int}}(C(s)) = \frac{1}{2}(\alpha|C'(s)|^2 + \beta|C''(s)|^2)$ is the internal energy with $\alpha$ and $\beta$ controlling the smoothness of $C$, and $E_{ext}(C(s))$ is the external energy, which attracts $C$ towards the object's boundary. $E_S$ can be minimized by evolving the snake dynamically as a function of parameter $s$ and artificial time $t$; namely, $C(s,t) = \left[\alpha C''(s,t) - \beta C''''(s,t)\right] - \nabla E_{ext}$, where the first term and the second term are called the internal force, $F_{\text{internal}}$, and the external force, $F_{external}$, respectively.

Our proposed external force is a weighted combination of balloon and GVF forces and aims at exploiting the advantages of each of these two forces. Balloon forces are represented as $F_{balloon} = n(s)$, where $n(s)$ is a unit vector normal to the snake at snake element $C_n$. GVF forces are derived from the diffusion of the gradient vectors of the image edge map $f(x,y)$, and are given as $F_{GVF} = \mathrm{v}(x,y)$, where $\mathrm{v}(x,y) = [u(x,y), v(x,y)]$ denotes the GVF field [11]. The proposed external force is then defined as $F_{external} = (F_{Balloon} * (1-\Omega)) + (F_{GVF} * \Omega)$, where $\Omega \in [0,1]$ is a weighting factor given by:

$$\Omega = \bar{h}^{(1-(\bar{AD}-\varepsilon))} \tag{1}$$

where $\bar{h} \in [0,1]$ denotes the average intensity value of $f(x,y)$ over a semi-circular region $S$ of radius $r$ centered at each snake element, and $\bar{AD} \in [0,1]$ is the average angular difference between the direction of $F_{Balloon}$ and the direction of the GVF field over a cone-shaped region $T$ centered at each snake element. $\bar{AD} = 0$ represents 0 rad, while $\bar{AD} = 1$ represents $\pi$ rad. Constant $\varepsilon = 0.001$ prevents power by zero when $\bar{AD} = 1$. For each snake element at position $(x,y)$, $\bar{h}$ and $\bar{AD}$ are calculated as follows:

$$\bar{h}(x,y) = \frac{1}{N}\sum_{(i,j)\in S} h(i,j); \quad \bar{AD}(x,y) = \frac{1}{M*\pi}\sum_{(i,j)\in T}\theta(i,j) \tag{2}$$

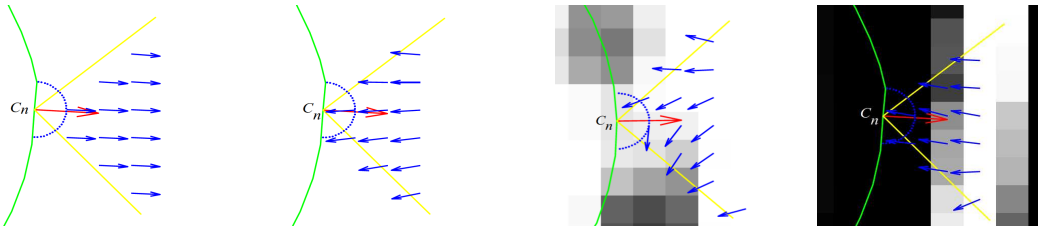
where $N$ is the number of pixels located in region $S$, $h$ is the edge intensity at position $(i,j)$, $M$ is the number of GVF field vectors in region $T$, and $\theta$ is the minimum angle between the GVF field vector at position $(i,j)$ and $F_{Balloon}$ for snake element $C_n$ at position $(x,y)$, as illustrated in Fig. 1. Note that $T(x,y,sl,\varphi)$ is a cone-shaped region with its vertex in $(x,y)$, which allows analyzing the region located far from snake element $C_n$ and thus provides a better insight of the direction of the GVF field than the region close to $C_n$.

(a) (b)

Figure 1: (a) Region $S$ and (b) $T$ for snake element $C_n$. The snake, the direction of $F_{Balloon}$, and the GVF field vectors are represented in green, red and blue, respectively.

The weighting function in Eq. (1) assigns different priorities to $F_{balloon}$ and $F_{GVF}$ according to two local image features: the average edge intensity ($\bar{h}$) and the average GVF field direction ($\bar{AD}$). This weighting process helps to achieve the following:

1. Guide the snake to the boundary in smooth areas even if its normal direction of growth is opposite to $F_{GVF}$. This allows initialization far from the boundary with a limited number of snake elements. Fig. 2(a-b) illustrate this, where $\Omega$ approaches 0.
2. Minimize snake leakages around weak edges by averaging the edge intensity and the GVF field direction. Fig. 2(c) illustrates this, where $\Omega$ slowly approaches 1.
3. Conform to the desired boundary by assigning a higher weight to $F_{GVF}$ when edges are encountered. This is illustrated in Fig. 2(d), where $\Omega$ approaches 1.



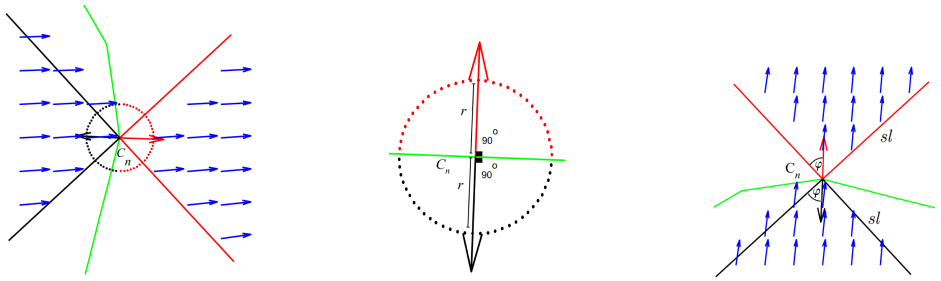

(a) (b) (c) (d)

Figure 2: GVF field direction in region $T$ (a) towards and (b) against the snake's growth ($\Omega \rightarrow 0$); and around (c) weak and (d) strong edges (non-white pixels represent strong edges).

In classical snakes, the deformation process is usually performed over a fixed number of iterations, over which it is expected to achieve convergence, i.e., external forces become zero. For images with strong edges, this number of iterations may be easily determined empirically. However, for images with weak edges, this number should be carefully selected to prevent leakages [7]. Here, we propose a mechanism to terminate the deformation process based on the percentage of snake elements labeled as *off*. A snake element is said to be *off* if no external force is acting on it. This usually occurs when the element encounters a strong edge, i.e., $F_{GVF}$ is the main acting external force ($\Omega > 0.5$) and the overall strength of the GVF field around the element is close to zero. To this end, if $\Omega > 0.5$, we compute $\Omega_b$ in the same manner as $\Omega$ in Eq. (1), but using regions $S_b$ and $T_b$, as illustrated in Fig. 3. We then label a snake element as *off* and set its forces to zero if $(\Omega_b)^2 > \Omega > 0.5$. This condition allows determining if the overall strength of the GVF field around the element is close to zero, i.e., the GVF field has opposite average directions in regions $T$ and $T_b$. We use the squared value of $\Omega_b$ to prevent the element from *moving* past the edge. If the percentage of *off* snake elements is equal to or greater than $Q$, the whole deformation process is terminated.

# 3 Experimental Results

The performance of the proposed approach over several regions of real MRI and CT slices is compared to that of snakes using three different external forces: a) $F_{GVF}$ exclusively [11], b) $F_{balloon}$ and image gradient (BGrad) forces [2], and c) $F_{GVF}$ or $F_{balloon}$ (BGVFT)




(a) (b) (c)

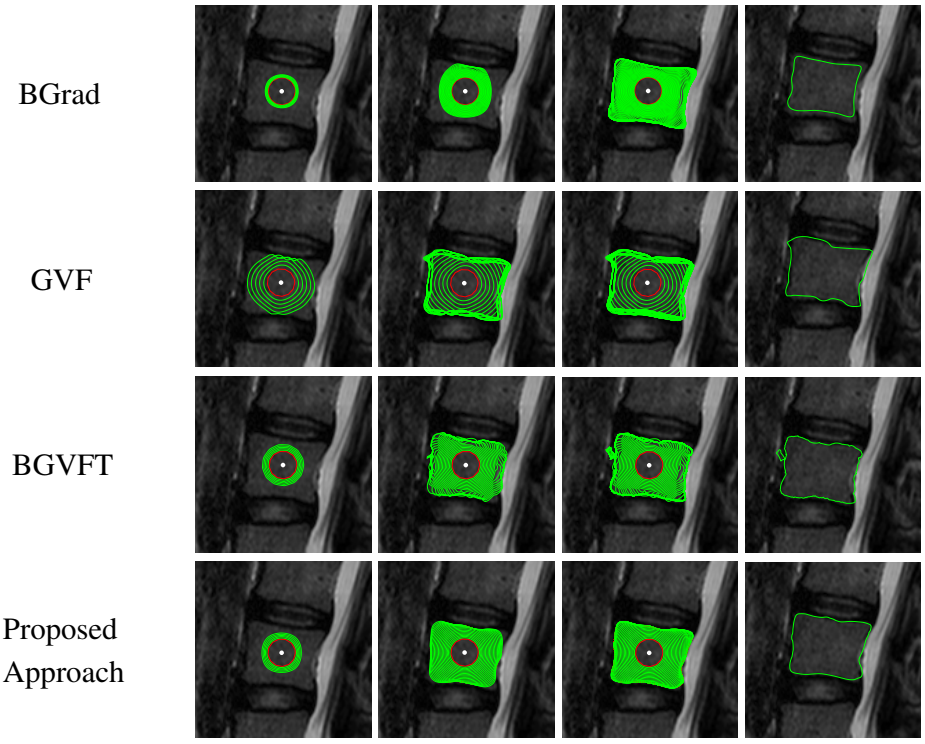
Figure 3: (a) Regions $T$ and $S$ (in red) used to compute $\Omega$, and regions $T_b$ and $S_b$ (in black) used to compute $\Omega_b$. (b) Dimension of regions $S$ and $S_b$ and of (c) regions $T$ and $T_b$.

Table 1: Detection accuracy of snakes using various external forces.

| Exp. | GVF | | BGrad | | BGVFT | | Proposed approach | | No. iterations |
|---|---|---|---|---|---|---|---|---|---|
| | DSC | JC | DSC | JC | DSC | JC | DSC | JC | |
| 1 | 0.8716 | 0.7724 | 0.8795 | 0.7848 | 0.9341 | 0.8763 | **0.9413** | **0.8902** | 43 |
| 2 | 0.9329 | 0.8742 | 0.8969 | 0.8130 | 0.9368 | 0.8812 | **0.9517** | **0.9079** | 45 |
| 3 | 0.9041 | 0.8249 | 0.8840 | 0.7921 | 0.9374 | 0.8822 | **0.9537** | **0.9115** | 44 |
| 4 | 0.9311 | 0.8710 | 0.8676 | 0.7662 | 0.9365 | 0.8806 | **0.9559** | **0.9154** | 38 |
| 5 | 0.9345 | 0.8770 | 0.8493 | 0.7380 | 0.9612 | 0.952 | **0.9631** | **0.9288** | 43 |
| 6 | 0.1445 | 0.0779 | 0.8017 | 0.6691 | 0.9517 | 0.9078 | **0.9643** | **0.9311** | 96 |
| 7 | 0.0044 | 0.0022 | 0.6877 | 0.5240 | 0.8996 | 0.8174 | **0.9709** | **0.9453** | 171 |
| 8 | 0.9381 | 0.8834 | 0.9045 | 0.8256 | **0.9568** | **0.9171** | 0.9433 | 0.8926 | 45 |
| 9 | 0.9043 | 0.8253 | 0.8263 | 0.7040 | **0.9393** | **0.8855** | 0.9310 | 0.8709 | 26 |
| 10 | 0.9258 | 0.8619 | 0.6844 | 0.5202 | 0.9231 | 0.8541 | **0.9426** | **0.8915** | 60 |

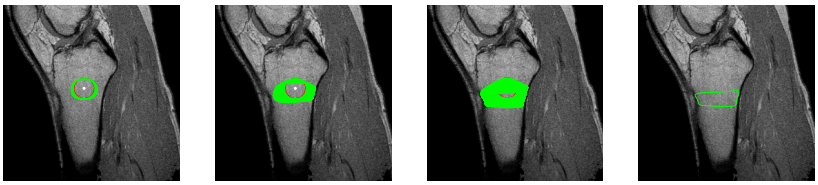
according to a threshold i.e., if $\bar{h} > 0.05$ the acting external force is then set to $F_{GVF}$. All images are preprocessed using histogram equalization to enhance edges. The edge map $f(x,y)$ is computed using the wavelet coefficients obtained after applying the stationary wavelet transform (SWT) using the Haar filter with $L = 3$ levels of decomposition, as explained in [9]. We use all parameter values suggested in [11] to compute the GVF field and to control the snake's smoothness. We use a radius $r = 1$ pixels for regions $S$ and $S_b$, an angle $\varphi = 45^o$ and $sl = 5$ pixels for regions $T$ and $T_b$. These values provide the best trade-off between capturing enough local image features and computational complexity. All snakes are initialized inside the desired region by manually selecting a single position that is used as the center of an initial circular curve of $p$ pixels in radius. The whole deformation process is terminated if more than $Q$ =90% of the snake elements are labeled as *off*.

The detection accuracy of the evaluated methods is measured by the Dice similarity (DSC) [3] and Jaccard coefficients (JC) [4] using manually annotated ground truth. DSC and JC values are within the range $[0,1]$, where 1 indicates identical overlap and 0 indicates no overlap between the regions inside the detected boundaries. Table 1 tabulates the DSC and JC values for different regions of MRI and CT slices. Experiments 1-3 represent three different regions of an MRI slice of a spinal cord, 4 and 5 represent two regions of an MRI slice of a pelvis, 6 and 7 represent two regions of an MRI slice of a knee, 8 represents one region of a CT slice of a skull, and 9 and 10 represent two regions of a CT slice of a spinal cord. In all experiments, the number of iterations is equal for all evaluated methods and is set to the number of iterations required by our termination mechanism. Results in Table 1 show that our approach achieves the highest accuracy for the majority of experiments. Note that for experiments 8-9, the BGVFT snake achieves higher DCS and JC values than our approach for the same number of iterations. In these cases, the BGVFT snake effectively switches between GVF and balloon forces based on the selected threshold. However, the




(a) 5 iterations (b) 20 iterations (c) 60 iterations (d) Final Results

Figure 4: Snake deformation process (green curves) for experiment 2. The white dot represents the manually selected initial position, the red curve represents the initial snake.



(a) 5 iterations (b) 20 iterations (c) 60 iterations (d) Final result

Figure 5: A GVF snake that fails if initialized by a small curve far from the desired boundary.

BGVFT snake, similarly to the GVF and BGrad snakes, requires manual selection of the number of iterations. More iterations may cause snake leakage, unlike our approach which terminates the deformation process when the majority of snake elements encounter strong edges. Visual results for experiment 2 are shown in Fig. 4. The BGrad snake conforms to the desired boundary with high accuracy (DSC=0.8969 ; JC=8130 ). However, for the same number of iterations required by our proposed approach, this snake does not reach the actual edge. More iterations may cause snake leakage around weak edges if the balloon force is greater than the gradient force. The GVF and BGVFT snakes result in leakages around weak edges. Our approach successfully conforms to the desired boundary with high accuracy.

Note that although the capture range of $F_{GVF}$ is in general large, methods using this external force exclusively may fail if the snake is initialized far from the desired boundary using a relatively small initial curve (see Fig. 5). The GVF field around such small initial snakes may not point towards the desired boundary due to the high levels of noise and clutter. An initial snake closer to the boundary may increase segmentation accuracy in this case.

# 4 Conclusion

This paper proposed an approach that combines balloon and GVF forces for parametric snakes, and a mechanism to automatically terminate the snake's deformation process. The

approach uses a weighting factor to leverage the advantages of these two forces according to local image features. In smooth areas, balloon forces guide the snake to the object's boundary, while in the presence of strong edge information GVF forces make the snake conform to the boundary. Experimental results on real medical images show that the proposed approach outperforms methods based on traditional external forces, while offering the advantage of initializing the snake with a single manually selected point inside the desired region and terminating the deformation process automatically.

# References


[1] Scott T Acton and Nilanjan Ray. Biomedical image analysis: segmentation. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 4(1):1–108, 2009.

[2] Laurent D Cohen. On active contour models and balloons. *CVGIP: Image understanding*, 53(2):211–218, 1991.

[3] Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[4] Paul Jaccard. The distribution of the flora in the alpine zone. *New phytologist*, 11(2): 37–50, 1912.

[5] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[6] Bing Li and Scott T Acton. Active contour external force using vector field convolution for image segmentation. *IEEE Trans. on Image Proc.*, 16(8):2096–2106, 2007.

[7] Liana M Lorigo, Olivier Faugeras, W Eric L Grimson, Renaud Keriven, and Ron Kikinis. Segmentation of bone in clinical knee MRI using texture-based geodesic active contours. In *Medical Image Computing and Computer-Assisted Interventation*, volume 1496, pages 1195–1204. Springer, 1998.

[8] Zhen Ma, João Manuel RS Tavares, Jorge R Natal, and T Mascarenhas. A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. *Comp. Methods in Biomechanics and Biomedical Eng.*, 13(2):235–246, 2010.

[9] G. Prince J. Clarkson S.-E.-A. Raza, V. Sanchez and N. M. Rajpoot. Registration of thermal and visible light images using silhouette extraction. *Pattern Recognition*, in press.

[10] Yuwei Wu, Yuanquan Wang, and Yunde Jia. Adaptive diffusion flow active contours for image segmentation. *Comput. Vis. Image Und.*, 117(10):1421–1435, 2013.

[11] Chenyang Xu and Jerry L Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369, 1998.

[12] Yuhua Yao, Lixiong Liu, Lejian Liao, Ming Wei, Jianping Guo, and Yinghui Li. Sigmoid gradient vector flow for medical image segmentation. In *11th International Conference on Signal Processing (ICSP)*, volume 2, pages 881–884. IEEE, 2012.

[13] Huiyu Zhou, Xuelong Li, Gerald Schaefer, M Emre Celebi, and Paul Miller. Mean shift based gradient vector flow for image segmentation. *Computer Vision and Image Understanding*, 117(9):1004–1016, 2013.

[14] Guopu Zhu, Shuqun Zhang, Qingshuang Zeng, and Changhong Wang. Gradient vector flow active contours with prior directional information. *Pattern Recognition Letters*, 31 (9):845–856, 2010.

[15] Christophe Zimmer, Elisabeth Labruyere, Vannary Meas-Yedid, Nancy Guillen, and J-C Olivo-Marin. Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: A tool for cell-based drug testing. *IEEE Trans. on Medical Imaging*, 21(10):1212–1221, 2002.

# Retinal Vessel Segmentation Using Gabor Filter and Textons

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**

This paper presents a retinal vessel segmentation method that is inspired by the human visual system and uses a Gabor filter bank. Machine learning is used to optimize the filter parameters for retinal vessel extraction. The filter responses are represented as textons and this allows the corresponding membership functions to be used as the framework for learning vessel and non-vessel classes. Then, vessel texton memberships are used to generate segmentation results. We evaluate our method using the publicly available DRIVE database. It achieves competitive performance (sensitivity (TPF) =0.7673, specificity (1-FPF) =0.9602, accuracy=0.9430), and especially its considerably higher sensitivity means it is better in detecting vessels than the other methods. These figures are particularly interesting as our filter bank is quite generic and only includes Gabor responses.

# 1 Introduction

Retinal vessel segmentation is an important preliminary stage in automatic assessment of retinopathy as it enables the vascular tree to be constructed. The integrity of the vascular tree is a key factor in assessment of systematic disease such as diabetes, glaucoma etc. Conventionally, the assessment of the retina vessel anomalies is a skilled time consuming task [1], and as such it has been the focus of research into automatic assessment techniques. Many techniques for automatic retinal vessel segmentation have been proposed [17][18][21][22][23] but a significant number use filter-banks to identify vessel and non-vessel structures. The success of filter-based methods depends largely on the design of the filter bank used to extract vessel features. The work of Chaudhiri et al. [2] is a good example of a filter-bank approach that employs a classic matched filter (MF) and several authors have built on this idea. For instance, M. Al-Rawi et al. [3] improved the performance of the matched filter by optimizing the filter parameters. In [4], Gang et al. proposed an amplitude-modified second-order Gaussian filter. They optimized the parameters of the matcher filter via mathematical analysis and experimental simulation. Zhang et al. [5] proposed a novel extension of the MF approach named MF-FDOG to distinguish the vessels from non-vessel step edges. Wu et al. [6] used a compound filter-bank that combines Hessian-based filters, matched filters and incorporates edge constraints

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

in a vessel enhancement stage. Further development of compound filters followed due to Adjeroh and Kandaswamy, et al. [7] (RVF-20) and Zhang and Fisher [9] (MR11) for retinal vessel extraction. However these approaches that rely on special filters sometimes introduce pathological changes in the fundus image causing these anomalies to be segmented as vessels, whilst some tinny vessels are miss-segmented. Although a compound vessel enhanced filter set can address some of these limitations, it is computationally expensive since many filter kernels need to be applied on the image. This work investigates if we can balance this situation by decreasing the number of filters while maintaining or improving the segmentation performance. Our technique uses the Gabor filter to enhance retinal vessels. The Gabor filter was originally proposed by Dennis Gabor [20] and subsequently used by Daugman [10][11] to model specific frequencies and orientations of certain cells in the visual cortex of some mammals. Because of its characteristics, Gabor filters have been widely used in many applications of image processing, such as object recognition, edge detection, and texture classification. In this paper, we present a procedure for parameter selection based on the retinal vessel features, and use a further parameter λ to control the function's performance. Section 2.1 describes details of this procedure. We optimise the parameters within the framework of textons derived from vessel feature vectors. Texton-based approaches have been a significant branch of texture analysis processes since the term texton was first introduced by Julesz in the 1980's [13]. Psychophysical experiments have shown that these texture primitives (microstructures) discriminate differences among various textures and models a process used by humans to discriminate one texture from another. Leung and Malik [14] described an operational definition of textons which enables textons to be generated automatically from an image [8][12].

We evaluate our experimental approach by training and testing on the publicly available DRIVE database [17]. The DRIVE data set contains 40 TIFF formatted RGB retinal images with a size of 565*584 pixels. The database is divided into training and test sets, each set consists of 20 images. For the test set, two sets of manual segmentation are provided by two observers. In practice, the first set is used as ground truth whilst the other one can be used to calculate the differences with the first one thus provides the measurements of independent human segmentation. The sensitivity (TPF), specificity (1-FPF) and the overall accuracy are used as performance measures of the experiments. Our results are compared with other state of the art approaches.

# 2 Method

In this experiment we choose a Gabor filter kernel to extract features of retinal vessels as the function encodes information about specific frequencies and orientations. The Gabor filter was originally introduced by Dennis Gabor [20] and defined as

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} e^{j\omega_0 x} \quad \text{(Eqn. 1)}$$

where $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$ represents a Gaussian envelope and $e^{j\omega_0 x}$ is a complex sinusoid. Daugman [11] extended the concept to two dimensions

$$g(x,y) = \bar{g}(x,y) e^{j[u_0(x-x_0)+v_0(y-y_0)]}$$
$$\text{where} \quad \bar{g}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{\frac{1}{2}\left(\frac{x^2}{\sigma_x}+\frac{y^2}{\sigma_y}\right)} \quad \text{(Eqn. 2)}$$

where $\sigma_x, \sigma_y$ determine the spread of the Gaussian envelope and $x_0, y_0$ the origin in the spatial domain. If the terms $x_0, y_0$ are set to 0, this Equation can also be written

$$G_{\lambda,\theta,\varphi,\sigma,\gamma}(x,y) = e^{\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)} \cos\left(2\pi\frac{x'}{\lambda}+\varphi\right) \quad \text{(Eqn. 3)}$$

where $x' = x\cos\theta + y\sin\theta$, $y' = -x\sin\theta + y\cos\theta$ and λ is the wave length (hence $1/\lambda$ is the spatial frequency) of the Gabor filter kernel. The angle parameter $\theta \in [0, \pi)$ determines the filter direction and specifies the orientation of normal parallel stripes of a Gabor function. The spatial aspect ratio γ determines the ellipticity of Gabor kernel (if γ equals 1, the kernel is circular) and σ is the standard deviation of Gaussian envelope.

## 2.1 Optimization of Gabor filter parameters

Because the vessel boundaries are normally presented and form (approximately) parallel edges, we model the vessel as an even symmetric function with $\varphi = 0$ and $\varphi = \pi$ corresponding to *center-on* and *centre-off* responses (Fig 1-c and 1-d). Since retinal vessels appear darker compared with their background (e.g. Fig.1-a) we choose $\varphi = \pi$ for our kernel model. Neurophysiological research shows that the parameter λ and σ are not independent [16]. Petkov and Kruizinga [15] reported that the ratio $\sigma/\lambda$ is related to the half-response spatial frequency bandwidth *b* and can be set as follows.

$$\frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{\ln 2}{2}}\cdot\frac{2^b+1}{2^b-1} \quad \text{(Eqn. 4)}$$

In practice, the bandwidth *b* controls the number of visible parallel excitatory and inhibitory stripe zones. Three zones, one inhibitory and two excitatory are visible in the retinal vessel structure (Fig.1-a) so we determined *b*=3 in our experiment. We set the spatial aspect ratio γ as 0.85 as our previous work suggests this ratio to be optimal. Since $\sigma$ and $\lambda$ are correlated, only one of them ($\lambda$) is considered a free parameter. Hence, we can write

$$G_{\lambda,\theta}(x,y) = e^{\left(-\frac{x'^2+0.72y'^2}{0.12\cdot\lambda^2}\right)} \cos\left(2\pi\frac{x'}{\lambda}+\pi\right) \quad \text{(Eqn. 5)}$$

Orientation of the retinal vessels is another significant structural characteristic since vessels are neither vertical nor horizontal precisely. Hence, the Gabor filter kernels are designed to cover 12 different orientations in 15° increments. An optimized kernel size is obtained by considering the value of σ. This plays a significant role in determining the width of vessels that will be detected. We choose $\lambda$ by plotting a family of ROC curves for a range of filter responses obtained from on a training set of images. Typical results are plotted in Figure 2 and from this analysis we chose the λ =13 (Fig. 1-e). Using Eqn. 4 we find σ is 3.12 and hence a suitable kernel size is 9 × 9.

## 2.2 Generating the Texton

A texton is an element (e.g. line segment, elongated blobs, cross and terminator) that is defined to represent the basic micro geometric textural structure in image [13]. The training procedure includes two stages. We trained vessel and non-vessel textons. In the vessel texton training stage our Gabor filter bank was applied to obtain filter response and the background was removed.

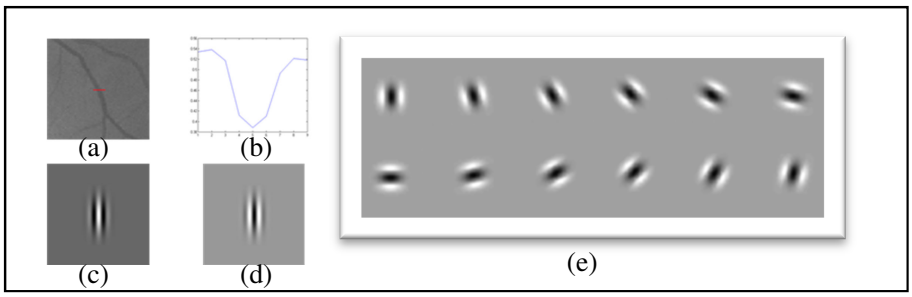



Figure 1: Showing the characteristics of vessel boundaries, the symmetric Gabor kernel and the Gabor filter bank. (a) is a panel cropped from a gray-level retinal image; (b) a gray level profile from the red line crossing the vessel in (a); (c) Gabor kernel with parameter φ=0; (d) Gabor kernel with parameter φ= π; (e) is optimized Gabor filter bank for retinal vessel feature extraction.

To train non-vessel textons vessel pixels were removed in order to obtain non-vessel influences background responses. The textons were generated by employing a *k*-means clustering algorithm on the filter responses. As representations of texture, the textons were aggregated based on the properties of distances calculated from memberships to clustering centres. This procedure is an iterative process.

Initially, K random points were selected as the means (centroids) of K clusters. Corresponding memberships were selected based on differences of Euclidean distance between means and centre bins. New means of those memberships were calculated again which were new clustering centroids. The process runs iteratively until it converges. Both vessel textons and background textons were stored in a texton dictionary and subsequently used in test stage. In test stage, the trained textons were assigned depending on responses of filter bank and corresponding memberships are calculated. The vessel texton relative memberships generate texton maps and the final segmentation results are obtained by combining various vessel texton relative maps. At textons training stage, various values of parameter K (no. of cluster centres) were chosen (K= 2,…,20). The corresponding accuracies were calculated by evaluating with respect to ground truth. The appropriate K=12 can be selected since the corresponding accuracy reaches maximum.

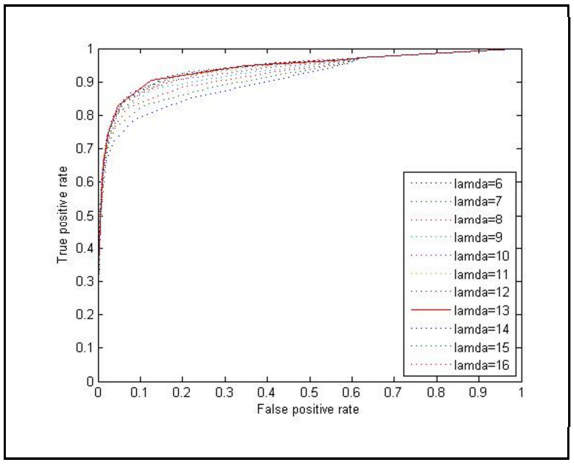



Figure 2: the ROC curves obtained by different lambda values.

# 3 Experiment results

The proposed method was tested and evaluated on DRIVE data sets. In order to quantify the performance of proposed approach, the resulting segmentation is compared to its corresponding ground truth. The ground truth is defined by a binary vessel mask in which all vessel pixels are set to one and all non-vessel pixels are set to zero. Our algorithm was evaluated in terms of sensitivity, specificity and accuracy for each sample of test images. The sensitivity (TPF) represents the true positive rate of segmentation and a significant factor in evaluating the performance of the method. The specificity (1-FPF) indicates the correct classification rate for segmentation of non-vessel pixels. The accuracy refers the overall segmentation performance. As shown in Table 1, for the DRIVE dataset, average specificity reaches 0.9602 with 0.7673 sensitivity, the average accuracy is 0.9430. To verify the performance of our proposed methods, we compare results with other state-of-the art approaches for retinal vessel segmentation. Table 1 presents corresponding results. The relative terms of measurement are average obtained from all of the test images. The experimental results show that our proposed method has produced a much better sensitivity, whilst maintaining almost the same overall accuracy, compared with the best other methods. In practice, it's difficult to balance the sensitivity and specificity. Normally, with increasing sensitivity, the value of specificity might reduce and this, in turn changes the overall accuracy. In our experiment, we find the sensitivity increase more than 5% however the specificity just reduce 1% compare with other methods. This verifies the robustness of our method in detecting retinal vessels.

| **Method** | **Performance Results** | | | |
|---|---|---|---|---|
| | ***database*** | ***Sensitivity*** | ***Specificity*** | ***Accuracy*** |
| 2nd observer | DRIVE | 0.7761 | 0.9725 | 0.9473 |
| Our method | DRIVE | **0.7673** | **0.9602** | **0.9430** |
| Mendonca [19] | DRIVE | 0.7344 | 0.9764 | 0.9452 |
| Zana[18] | DRIVE | 0.6696 | 0.9769 | 0.9377 |
| Staal [17] | DRIVE | 0.7194 | 0.9773 | 0.9441 |
| Zhang [5] | DRIVE | 0.7120 | 0.9724 | 0.9382 |
| Soares [22] | DRIVE | 0.7283 | 0.9788 | 0.9466 |
| Fraz[23] | DRIVE | 0.7525 | 0.9722 | **0.9476** |

Table 1: comparative results on DRIVE database

# 4 Conclusion

We proposed a novel method for retinal vessel segmentation by using both Gabor filter and texton. A new filter bank for retinal vessel network enhancement was designed based on the human visual system and the corresponding parameters of the Gabor filter were optimized using a ROC analysis of the filter performance on training data. Our experimental results demonstrated that our method significantly enhances the true positive rate while maintaining a level of specificity that is comparable with other approaches. The overall performance accuracy (0.943) is better than most of state-of-the-art retinal vessel segmentation methods proposed by other authors. Although the performance of our method in terms of sensitivity, specificity, and accuracy is good, we can find some limitations and weaknesses from our method. Visually, some false positive pixels are evident in our segmentation results, especially in an area around the optic disc (OD). This is because the area surrounding the OD exhibits strong contrast and so there are significant gradient changes on its boundary. We intend to address these limitations in our further work.

# References


[1] H. Schneiderman, The Funduscopic Examination, 1990, Website access at http://www.ncbi.nlm.nih.gov/books/NBK221/.

[2] S. Chaudhuri et al., Detection of blood vessels in retinal images using two-dimensional matched filters. IEEE Tran. on Medical Imaging, 8(3):263 269, September 1989.

[3] M. Al-Rawi, M. Qutaishat, M. Arrar, An improvement matched filter for blood vessel detecetion of digital retinal images, *Comput. Biol. Med.* 37:262–267, 2007.

[4] 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., 49(2):168–172, 2002.

[5] 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

[6] C. H. Wu, G. Agam, P. Stanchev. A general framework for vessel segmentation in retinal images, in Proceedings of CIRA. 37-42, 2007.

[7] D. A. Adjeroh, U. Kandaswamy, J. V. Odom, Texton-based segmentation of retinal vessels. *Journal of the Optical Society of America*, 24(5): 1384-1393, 2007.

[8] M. Varma and A. Zisserman, A statistical approach to texture classification from single images," Int. J. Comput. Vis., 62: 61–81, 2005.

[9] M. Fisher, L. Zhang, W. J. Wang, Locating blood vessels in retinal images using unified Textons, MIUA 2013, 213-218, 2013.

[10] J. G. Daugman. Two dimensional analysis of cortical receptive field profiles. *Visual Research*, 20, pp. 846-856, 1980.

[11] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America*, 2(7):1160–1169, July 1985.

[12] M. Varma and A. Zisserman. A Statistical Approach to Texture Classification from Single Images. *International Journal of Computer Vision*. 62:61-81.2005.

[13] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802):91–97. 1981

[14] T. Leung and J. Malik, Representing and recognizing the visual appearance of materials using three dimensional textons, *Int. J. Comput.* 43:29–44, 2001.

[15] N. Petkov, P. Kruizinga, Computational models of visual neurons specialized in the detection of periodic and aperiodic oriented visual stimuli: Bar and grating cells, *Biol. Cybern.*, 76:83–96, 1997.

[16] N. Petkov, Biologically motivated computationally intensive approaches to image pattern recognition, *Future Gen. Comput. Syst.*, (11):451–465, 1995.

[17] J. J. Staal, et al., Ridge based vessel segmentation in color images of the retina, IEEE Trans. Med. Imaging, 501–509, 2004.

[18] F. Zana and J.-C. Klein, Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation, IEEE Trans. Image Process. 10:1010–1019, 2001.

[19] A. M. Mendonca, A. Campiho, Segmentation of retinal blood vessels by combining the detection of centrlines and morphological reconstruction, IEEE Trans, Med, Imaging 25(9): 1200-1213,2006.

[20] D. Gabor, Theory of communication. *Journal of the Institute of Electrical Engineers,* 93:429-457, 1946.

[21] A. Hoover, V. Kouznetsova, M. Goldbaum, Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, IEEE Trans. Med. Imaging, 19(3): 203–210, 2000.

[22] J. V. B. Soares, et al., Retinal vessel segmentation using the 2D Gabor wavelet and supervised classification, IEEE Transactions on Medical Imaging, 25(9): 1214–1222, 2006.

[23] M.M. Fraz, P. Remagnino, et al., A Supervised Method for Retinal Blood Vessel Segmentation Using Line Strength, Multiscale Gabor and Morphological Features, in IEEE Inter. Conf. on ICSIPA, 410-415, 2011.

# Computer-Aided Segmentation and Estimation of Indices in Brain CT Scans

Adnan Qureshi
adnan.qureshi@beds.ac.uk
Vitaly Schetinin
vitaly.schetinin@beds.ac.uk

Institute for Research in Applicable Computing
University of Bedfordshire
Luton, UK.

**Abstract**

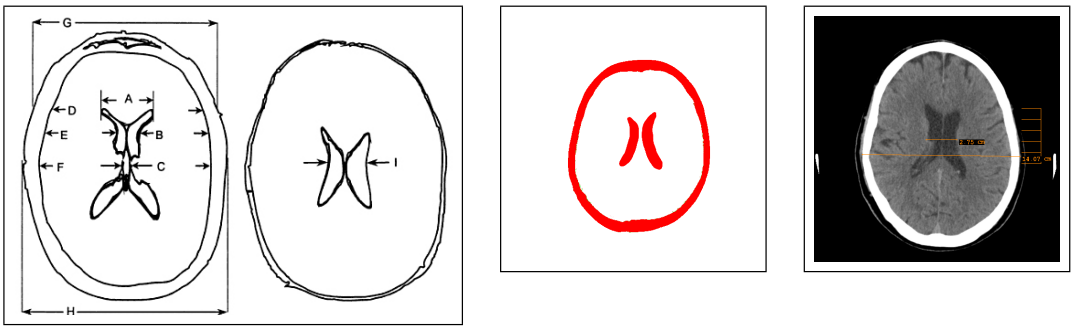
The importance of neuro-imaging as one of the biomarkers for diagnosis and prognosis of pathologies and traumatic cases is well established. Doctors routinely perform linear measurements on neuro-images to ascertain severity and extent of the pathology or trauma from significant anatomical changes. However, it is a tedious and time consuming process and manually assessing and reporting on large volume of data is fraught with errors and variation. In this paper we present a novel technique for segmentation of significant anatomical landmarks using artificial neural networks and estimation of various ratios and indices performed on brain CT scans. The proposed method is efficient and robust in detecting and measuring sizes of anatomical structures on noncontrast CT scans and has been evaluated on images from subjects with ages between 5 to 85 years. Results show that our method has average ICC of $\geq 0.97$ and, hence, can be used in processing data for further use in research and clinical environment.

## 1 Overview and problem statement

Linear measurements on axial CT scans provide clinicians and surgeons opportunity to ascertain differential diagnoses of neuropsychiatric disorders, outcomes of clinical and surgical interventions, geriatric changes and deleterious effects of drugs. Quantitative assessment of neuro-images is an effective approach to reveal structural changes in conditions such as Alzheimer's disease (AD), Schizophrenia, Huntington's disease, hydrocephalus and many other neurological and psychiatric disorders [1, 2, 11, 12]. The typical measurements performed on the axial CT scans (fig. 1a, details in table 1) are used to estimate indices and ratios such as Evan's Ratio (ER), Bifrontal Index (BFI), Bicaudate Index (BCI), Cella Media Index (CMI), Frontal Horn Index (FHI), Ventricular index (VI), Huckman number (HN) and 3rd ventricle width (V3) [3]. Manual measurements of regions of interest (ROI) are still considered 'gold standard', however, these are time consuming and a robust and efficient method is required to assist the researchers and to our knowledge, there is no automated system for performing measurements and calculation of these from CT scans.

Computer aided quantitative radiology and volumetry studies of the human anatomy and pathology have been undertaken by many researchers and a surfeit of image segmentation methods have been proposed [10]. However, in medical images an ROI can have dissimilarity of pixel intensity, inhomogeniety of background contribution and noise, and spatially

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.




(a) (b) (c)

Figure 1: Measurements on Axial CT Scans (a) and Manual Labelling (b, c).

blind methods in these cases can result in disjoint, quasi-homogeneous regions [9]. To preserve the spatial relationship between the pixels and neighbourhood information, spatially guided techniques such as region growing and merging, active contours and level sets show better results [10].

The extensive and tenable integration of computational intelligence in medical problems can be attributed to the fact that these systems can adaptively learn and optimize the relationship between inputs and outputs [4, 5, 6, 7, 8]. Medical images e.g., CT scans, usually have inhomogeniety of background and noise; and recognition of ROI in images with similar characteristics using ANN can give plausible results [13]. In addition, ANN's can be trained using a few images which have been manually annotated and labelled in order to learn to recognize the ROI.

# 2 Methodology

The proposed methodology uses ANN to segment the input images into ROI. Measurements of landmark regions are then used for calculating the various ratios and indices. The input images are DICOM (Digital Imaging and Communications in Medicine) files and the training set has been manually segmented and measured by radiology experts (Figure 1b, 1c). The dataset used for training, validation and testing consists of noncontrast CT studies from different subjects. The studies include subjects of both genders and between ages 5 to 85 years. Our method is used on axial CT slices.

## 2.1 Correction of Orientation of Head

The first step is to align the head into proper orientation on the CT image to ensure that the reference points for measurements are properly identified. This is done by demarcating the bony protuberance of the anterior and posterior falx cerebri attachments as these landmarks are usually not affected in traumatic and pathological cases in which the brain tissue may show abnormalities. The skull boundaries are extracted using Canny edge detection with $\sigma = 2$ and $threshold = 0.5$ in our experiments, and properties of connected regions are measured. Then the curve of the skull bone is modelled as a function of the location and height of pixels and the extrema represent the protuberances which are used to find any shift $P_{shift}$ in the orientation.

(a) (b) (c)

Figure 2: Correction of orientation of CT Scans.

Taking *arctan* of the $P_{shift}$ with reference to the skull centroid calculated earlier, gives the rotation angle $\theta^{\circ}$ to be corrected (Figure 2). An ideal midline can also be drawn using skull centroid which, in normal cases, would pass through the septum pellucidum. Identifying the septum pellucidum and its location with reference to skull can be used for measuring the midline shift which is an important clinical feature in assessment of traumatic or pathological conditions.

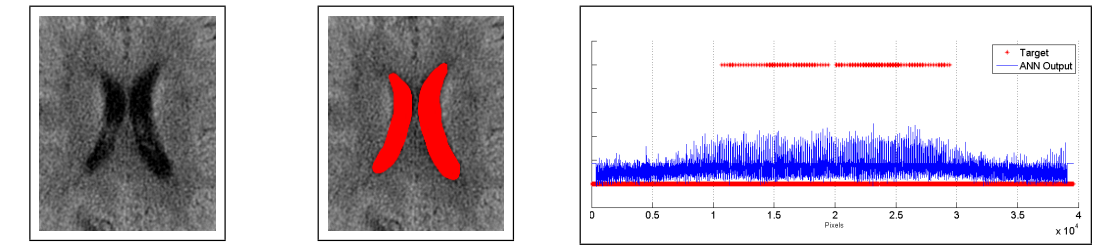
## 2.2 Detection of Candidates for Measurements

An input image is represented as an $(n \times m)$ matrix $X = \{x_{ij}\}, i = 1 \ldots n, j = 1 \ldots m$, whose pixels can either belong to a region of interest class $\Omega_0$ or to a non region $\Omega_1$. A function $u_{ij} = \phi(X; i, j)$ reflects the brightness of the pixel $x_{ij}$ influenced constantly or variably by the neighbourhood. The function is estimated in a rectangular window $P$, which is a $(k \times k)$ matrix containing the central pixel $x_{ij}$ and $r-1$ nearest pixels, where $r = k^2$. Then by sliding the window $P$ through the image $X$, every central pixel $x_{ij}$ can be assigned to one of the two classes $\Omega_0$ or $\Omega_1$. The sliding window transforms the image $X$ into an $r \times q$ matrix $Z$ where $q = (n-k+1)(m-k+1)$ and the central element of a vector $z \in Z$, hence, represents the pixel $x_{ij}$ of the input image. The matrix $Z$ is fed to the ANN with one hidden layer and the output neuron makes a decision $y_c = \{0,1\}, c = 1 \ldots q$, on the central element of $c$th column vector $z^{(c)}$, thus, classifying the image pixels to classes $\Omega_0$ or $\Omega_1$ as proposed by [13]. The ANN is trained with standard back propagation technique aiming to minimize an error function $\varepsilon$ for a given number of training epochs. We use images whose pixels are manually labelled and assigned either to class $\Omega_0$ or $\Omega_1$ by a radiology expert Fig. 3b and the ANN is trained (Fig. 3c) to approximate these. The approach is implemented to identify both ventricles and skull as ROI for subsequent measurements and estimations of indices.

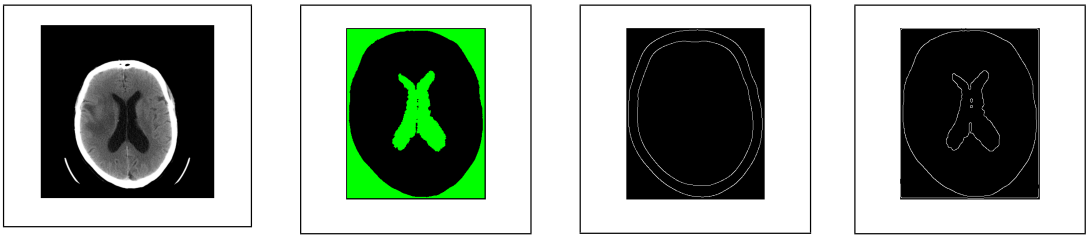
The output from the neural network classifies pixels of the input image into ROI as shown in Fig. 3e. In this binarized image, the discontinuities within the regions and the non-smooth boundaries are further refined using active contours using output of ANN for automatic seeding. The values of curve length $\mu$ and curvature $\kappa$ are empirically estimated in our experiments to ensure a good approximation of the edges of regions. This step generates a binary image as output with the skull and ventricles' boundaries clearly demarcated as in Fig. 3f and 3g respectively.

## 2.3 Measurements and Estimation of Indices

The CT scan studies are viewed using 'brain window' (width = 80, length = 40) with image size of $512 \times 512$ pixels and a reference ruler is overlaid on the images to calibrate our




(a) Input image (b) Manual labelling (c) ANN Output



(d) Input Image (e) Segmentation (f) Skull (g) Ventricles

Figure 3: Segmentation of the ROI in Brain CT Scans.

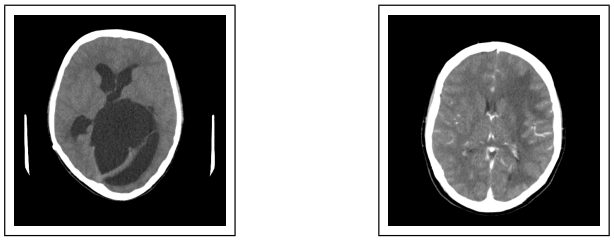
measurement algorithm. An expert then provides multiple reference measurements in an interactive session using a simple user interface and the averages are stored as benchmarks $P_{cal}$ for subsequent automatic measurements. There is an option to use Metric or Imperial units for measurements. The contiguous image regions from step 2.2 are used to determine measurements such as extrema, centroid, area, major and minor axes. The values of extrema are used to draw reference points (Fig. 1a) around the ROI using same approach as section 2.1 and the number of pixels between relevant points is divided by $P_{cal}$ to give measurements. To estimate the indices, the distance between relevant reference points is counted in terms of pixels and appropriate formula is applied from atlas [3].

# 3 Experimentation Results

The technique has been applied on 48 noncontrast CT studies from different subjects with ages between 5 to 85 years including 14 subjects from the National Cancer Institute public dataset. Two radiology experts independently performed manual segmentation and calculated ratios and indices to be used for training and evaluation in our experiments. The subjects' personal information was anonymized and only age and gender data were retained. Measurements of the maximum frontal horn diameter (A), inner skull diameter at the level of frontal horns (D), maximum inner skull diameter (F), outer skull diameter at the frontal horns (G), maximum outer skull diameter (H) and minimum width of the later ventricles separated by the septum (I) were made by experts. Our method repeated the same linear measurements after segmenting the ROI and using reference points and calibration information. The results were compared with the measures in noncontrast CT scans from different age groups and showed ICC $\geq 0.97$. Correction of orientation of head showed 85% sensitivity and 75%, while detection of brain midline shift $> 5mm$ showed sensitivity and specificity of 80% and 86% respectively. The intraclass correlation coefficient (ICC) was calculated using MedCalc to compare the results with experts to ascertain reliability of measurements. Linear measurements show ICC range from 0.97 to 0.99 and the results are given in table 1. For both the CMI and ER estimations, ICC was 0.98 compared to the experts. The optimum testing performance of 88.9% of the ANN was observed with 10 hidden neurons to segment the

| Label | Measurement | ICC |
|---|---|---|
| A | Maximum frontal horn diameter | 0.99 |
| D | Inner skull diameter at the level of the frontal horns | 0.97 |
| F | Maximum inner skull diameter | 0.97 |
| G | Outer skull diameter at the frontal horns | 0.97 |
| H | Maximum outer skull diameter | 0.98 |
| I | Minimum width of the lateral ventricles separated by the septum (cella media) | 0.98 |

Table 1: Measurements on the axial CT scans



(a) Transtentorial Arachnoid cyst (b) Metastatic lesions

Figure 4: Cases excluded from experiments.

image using 70% training, 15% validation and 15% test samples. Increasing or decreasing the number of hidden neurons in ANN did not significantly improve the performance and the average performance was $87.4-88.9\%$ in training and testing. The sliding window in section 2.2 was evaluated with varying sizes of $3\times3$, $5\times5$ and $7\times7$ neighbourhood pixels, however, the optimum output was achieved with $5\times5$ pixels windows for extracting features. The experiments were performed using Matlab R2013b on Mac OS X and Windows 7 platforms. The average time taken by our method to perform segmentation and estimations of indices and ratios was $\leq 20$ seconds because the algorithm performs the segmentation and measurements in toto without performance degradation. CT studies of patients with gross deformity of the brain structure due to either neoplastic metastasis, pathology or trauma gave less satisfactory estimations and resulting measurements were amiss. These 2 of the 34 cases (fig. 4) were excluded from the performance analyses after preliminary experiments.

# 4 Conclusion

Manual demarcation of ROI by expert radiologists is still considered gold standard, however, our proposed system has shown plausible results with $87.4-88.9\%$ accuracy with ANN in segmenting the ROI which are used as reference points for estimating various brain indices. The error in linear measurements on noncontrast axial CT scans was less than $\pm1mm$ and ICC results were $\geq 0.97$ compared with experts. Correction of orientation of head and detection of midline shift $> 5mm$ showed sensitivity of 85% and 80% and specificity of 75% and 86% respectively. The average time taken for measurements was also significantly reduced. Currently, the CT slice with pertinent anatomical landmarks is manually input to the system for measurements. However, we are working on an algorithm to automatically select the representative slice from the complete study of the subject. Interpretation of the

medical images by the radiologists is fraught with errors and variations which represent the weakest aspect of clinical imaging. Based on the results of our method, we propose that integration of automated systems in clinical and research settings can significantly reduce the radiologists' workload and inter-observer variability in assessing normal development, ageing, pathological and traumatic cases.

# References

[1] S. Barker-Collo, M. Kahan, and N. Starkey. Computerised tomography indices of raised intracranial pressure and traumatic brain injury severity in a new zealand sample. *The New Zealand medical journal*, 125(1360):92, 2012.

[2] E. González-Reimers and F. Santolaria-Fernández. Brain atrophy in alcoholics. In *Handbook of Behavior, Food and Nutrition*, pages 2993–3010. Springer, 2011.

[3] T. Keats and C. Sistrom. *Atlas of radiologic measurement*. Mosby Inc, 2001.

[4] T. Kondo and J. Ueno. Feedback GMDH-type neural network and its application to medical image analysis of the liver cancer. In *42th ISCIE international symposium on stochastic systems theory and its applications*, pages 81–82, 2012.

[5] I. Kononenko. Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine*, 23(1):89–109, 2001.

[6] V. Schetinin and J. Schult. A neural-network technique to learn concepts from electroencephalograms. *Theory in Biosciences*, 124(1):41–53, 2005.

[7] V. Schetinin and J. Schult. Learning polynomial networks for classification of clinical electroencephalograms. *Soft Computing*, 10(4):397–403, 2006.

[8] Z. Shi and L. He. Application of neural networks in medical image processing. In *Proceedings of the Second International Symposium on Networking and Network Security*, pages 2–4, 2010.

[9] W. Tao, H. Jin, and Y. Zhang. Color image segmentation based on mean shift and normalized cuts. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(5):1382–1389, 2007.

[10] S. Vantaram and E. Saber. Survey of contemporary trends in color image segmentation. *Journal of Electronic Imaging*, 21(4):040901–1, 2012.

[11] R. Wilk, E. Kluczewska, B. Syc, and G. Bajor. Normative values for selected linear indices of the intracranial fluid spaces based on ct images of the head in children. *Polish Journal of Radiology*, 76(3):16, 2011.

[12] Y. Zhang, E. Londos, L. Minthon, C. Wattmo, H. Liu, P. Aspelin, and L. Wahlund. Usefulness of computed tomography linear measurements in diagnosing alzheimer's disease. *Acta Radiologica*, 49(1):91–97, 2008.

[13] V. Zharkova and V. Schetinin. Filament recognition in solar images with the neural network technique. *Solar Physics*, 228(1-2):137–148, 2005.

# On Bloodvessel Branching Analysis for the Detection of Alzheimer's Disease

Musab Sahrim[1]
m.sahrim@soton.ac.uk
Mark S. Nixon[1]
msn@ecs.soton.ac.uk
Roxana Carare[2]
R.O.Carare@soton.ac.uk

[1] School of Electronics and Computer Science
University of Southampton UK
[2] School of Medicine
University of Southampton UK

**Abstract**

Alzheimer's Disease (AD) is increasingly prevalent in modern society and methods for its diagnosis are only just starting to emerge. Given images of brain tissue, we show how Alzheimer's disease can be detected from the branching structures of blood vessels. This is achieved by a new approach which counts the branching points and derives measures which are suited to the analysis of small branching structures. The measures are formulated to be rotation, scale and position invariant and are deployed in tandem with more standard measures. Analysis on a database comprised of brain tissue samples from subjects who are normal, with Alzheimer's and age matched normal has shown capability to classify correctly images of brain tissue from subjects afflicted with Alzheimer's disease.

# 1 Introduction

Alzheimer's Disease (AD) is the commonest form of dementia, affecting nearly a million people in the UK and with no effective treatment. Diagnosis is difficult, as many neurodegenerative conditions present similar characteristics [1]. Amyloid beta (Aβ) is a normal product of metabolism, cleaved from an amyloid precursor protein (APP). Young brains are equipped with different mechanisms to break down and eliminate Aβ, but with ageing and on the background of different genotypes the elimination of Aβ fails, leading to its accumulation and to AD [2]. The accumulation of Aβ in the walls of blood vessels of the brain reflects a failure of its elimination along the walls of blood vessels.

In recent years, researchers have tried to detect AD by processing images of the human brain. Most have used MRI and CT scans to detect the abnormalities. For example, shape changes were detected in the corpus callosum in AD [3]. In addition, texture was evaluated to discriminate a brain with AD from one without [4], while a new method was introduced to measure the thickness of the human cerebral cortex by considering the white and the grey matter. However although the methods using computer vision have been demonstrated some detection capability, little attention has been given to detecting the abnormalities of specific components in the brain that are affected by amyloid beta, such as the blood vessels. The concept of the early onset detection of AD has yet to receive much

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

research attention. Naturally, any approach that can detect AD at the onset or early in its progression could be invaluable in medical planning.

Blood vessels have previously been analysed in diagnosis of diabetes, hypertension and atherosclerosis [5] supporting their potential use for Alzheimer's Disease diagnosis. The previous importance of the blood vessels suggests detecting abnormalities of the blood vessels is important in human condition analysis, and coupled with observations concerning Aβ drainage suggests that analysis of the branching structure might give an indication of the presence of AD. There is a general computer vision aspect in branching in that natural patterns such as sea fans, water flow in the rivers and trees show structural similarity suggesting automated analysis is appropriate there as well.

The detection of a blood vessel's pattern can use many established features such as density and tortuosity for Alzheimer's disease detection [6]. In this study, we describe the blood vessel's overall structure of by their density, by a new invariant measure of branching structure, and by their tortuosity. Our aim is to detect the signs of AD at an early stage by focusing on the objects that have been most affected by this disease.

Current approaches to analysing branching structures in medical images currently concentrate on ductal trees [7] and vascular structures [8]. These are largely large scale analyses of topology whereas blood vessels in segments of brain tissue are considerably smaller. Further, there is no analysis of this sort directed towards the detection of Alzheimer's disease.

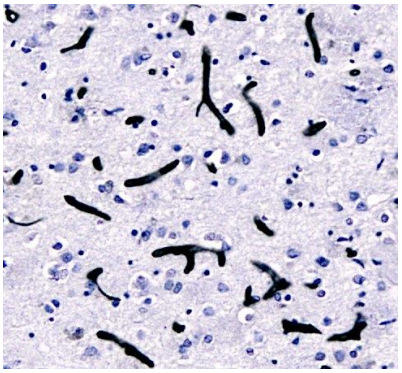


Figure 1: Example Image of Brain Tissue

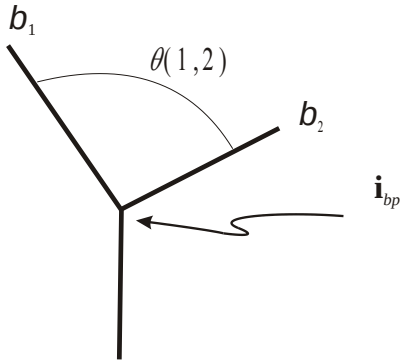



Figure 2: Branching Structure

In this paper, we propose a new analysis of the branching structure, with a measure which is invariant to rotation, scale and translation. This measure together with density and tortuosity are analysed on a database of images of brain tissue derived from young people, old people with Alzheimer's disease and age-matched old people without Alzheimer's disease to show that these new measures have detection capability.

# 2 Materials and Methods

## 2.1 Brain Images

We have used images of human tissue from the Brain Tissue Resource in Newcastle derived using a Nikon Eclipse E600 microscope fitted with a digital camera Nikon Coolpix 950 at fixed ×10 magnification, Fig. 1. We have five samples of tissue from; young, older

age-matched control and older Alzheimer's subjects. This dataset appears small compared with large scale image analysis datasets, and the purpose here is proof-of-concept and this size suffices for that purpose.

## 2.2 Density

The density is defined as the number of branches in the image which is determined by detecting the points at which branching occurs. Images are first thresholded using Otsu's technique and then thinned to provide a skeleton. A potential branching point $I(V)$, is calculated for each point $V$, of the skeleton where the $N_i(V)$ are the neighbours $i$ of the analysed point, $V$. Should there be more than two neighbours $V$ is classified as a branching point $\boldsymbol{i}_{bp}$, Eqs. 1 &2.

$$I(V) = \frac{1}{2}\left(\sum_{i=1}^{8} |N_i(V) - N_{i+1}(V)|\right) \tag{1}$$

$$\boldsymbol{i}_{bp} = I(V) > 2 \tag{2}$$

After obtaining the branching points $\boldsymbol{i}_{bp} = (x_{bp}, y_{bp})$, the branching points are then excluded when counting the number of branches, giving the density $D$. The vessels $I_{vessel}$ are determined via 8-way connectivity and counted excluding the branching points as

$$D = \sum \left(I_{vessel} \cup \overline{I_{bp}}\right) \tag{3}$$

## 2.3 Description of the Branching Structure

The detected branching points allow for the new analysis of the branching structure. Essentially we have segments of blood vessels which are at different inclinations to the branching points. For a branching point $bp$ with $N$ branches each of length $L(n)$ the branching structure, as shown in Fig. 2, can be described by a composite measure which is derived from the branch length and relative inclination of the vessels. The average vector product of pairs of branches and the angle between them, gives the measure $B$ which is invariant to rotation, translation and scale (size).

$$B = \sum_{i=1,N,j=1,N} L(i) \times L(j) \times \cos(\theta(i,j)) / N \quad i \neq j \tag{4}$$

As this equation appears to favour smaller angles, a version not using the cosine was also deployed, though this was found to have lower performance.

## 2.4 Other measures

We also described vessel structure using a standard measure of tortuosity $\tau$

$$\tau = c / L \tag{5}$$

where $c$ is the total length of the curve and $L$ is the length of a straight line connecting the curve end points. This measure suffices for the analysis here though could be extended [9]. Here the tortuosity is averaged over all detected branches in an image.

# 3 Results

Figure 3 shows the results of correct classification rate achieved by using the function shown in eq. 4, and classified using the *k* nearest neighbour rule across the whole dataset and compared to version that used the pure angle (without the cosine function). Clearly, using the averaged dot product improves the correct classification rate.

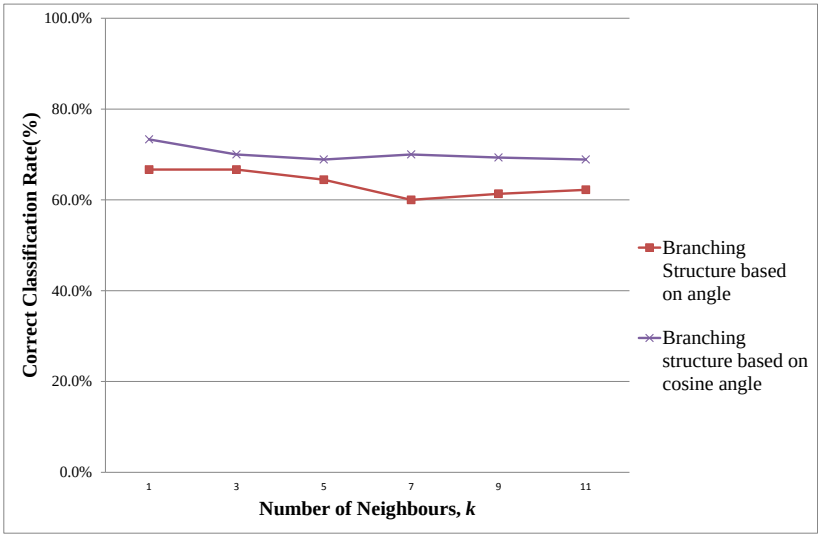
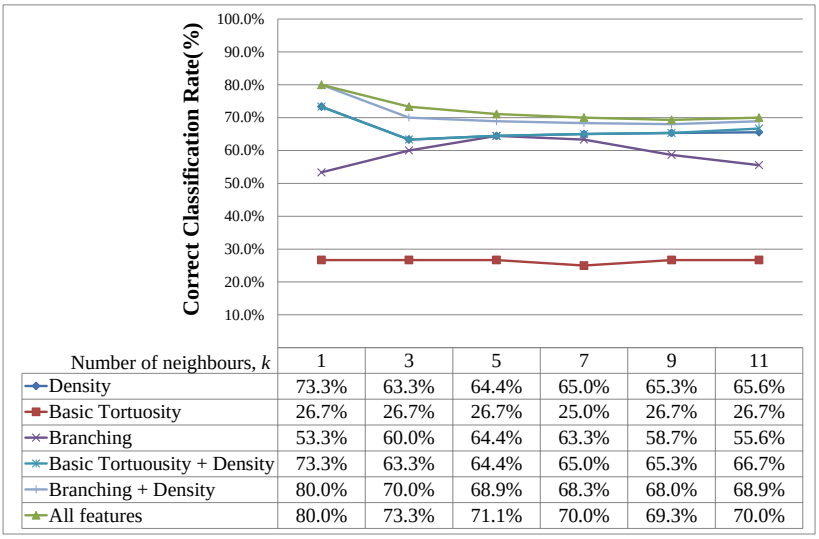



Figure 3: Correct classification rate using alternative branching measures.

Next, Figure 4 shows the Correct Classification Rate for the images when divided into three groups; young, old-matched control and old with Alzheimer's disease. Here, samples are classified by correct classification is achieved when an image from a sample without Alzheimer's disease is correctly classified as not having the disease. Correct classification also occurs when a sample from an older subject with Alzheimer's disease is correctly classified as that. The horizontal axis is the value of *k* used within the *k*-NN rule.

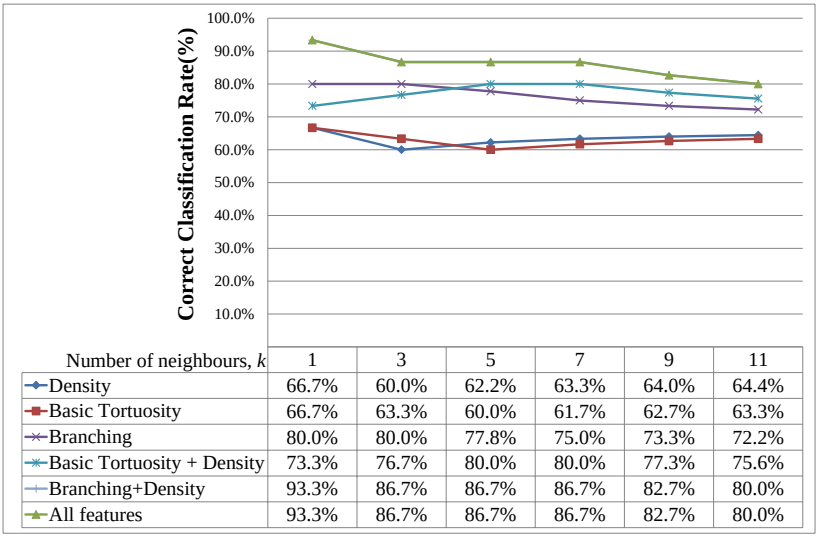
In Fig. 4 the features used are the density $D$, the tortuosity $\tau$ or the branching measure $B$ or combinations thereof, combined as a feature vector for use with the *k*-NN rule. By this analysis, each measure contributes to successful discrimination of subjects with Alzheimer's disease, suggesting that image based analysis of brain tissue samples could prove a suitable avenue for research, as predicted by physical analysis and the postulated effects of Aβ.




| Number of neighbours, *k* | 1 | 3 | 5 | 7 | 9 | 11 |
|---|---|---|---|---|---|---|
| Density | 73.3% | 63.3% | 64.4% | 65.0% | 65.3% | 65.6% |
| Basic Tortuosity | 26.7% | 26.7% | 26.7% | 25.0% | 26.7% | 26.7% |
| Branching | 53.3% | 60.0% | 64.4% | 63.3% | 58.7% | 55.6% |
| Basic Tortuousity + Density | 73.3% | 63.3% | 64.4% | 65.0% | 65.3% | 66.7% |
| Branching + Density | 80.0% | 70.0% | 68.9% | 68.3% | 68.0% | 68.9% |
| All features | 80.0% | 73.3% | 71.1% | 70.0% | 69.3% | 70.0% |

Figure 4: Correct classification rate and for each feature combinations versus the number of nearest neighbours *k*

Next, in Figure 5 we combine the young datasets with the control dataset to perform the discrimination of the brain tissue with AD from normal brain tissue. This provides ten images from young and age-matched control subjects, together with five samples from subjects with Alzheimer's Disease.




| Number of neighbours, *k* | 1 | 3 | 5 | 7 | 9 | 11 |
|---|---|---|---|---|---|---|
| Density | 66.7% | 60.0% | 62.2% | 63.3% | 64.0% | 64.4% |
| Basic Tortuosity | 66.7% | 63.3% | 60.0% | 61.7% | 62.7% | 63.3% |
| Branching | 80.0% | 80.0% | 77.8% | 75.0% | 73.3% | 72.2% |
| Basic Tortuosity + Density | 73.3% | 76.7% | 80.0% | 80.0% | 77.3% | 75.6% |
| Branching+Density | 93.3% | 86.7% | 86.7% | 86.7% | 82.7% | 80.0% |
| All features | 93.3% | 86.7% | 86.7% | 86.7% | 82.7% | 80.0% |

Figure 5: Correct classification rate overall and for each feature combinations versus the number of *k* using 2 groups

Clearly in this figure, the CCR for branching structure shows capability to differentiate the AD from other normal brain. By combining the density with branching structure, we can see some enrichment of CCR suggesting the continuity of this combination for this research.

# 4 Conclusions

Studies suggest that the drainage of the protein Aβ is consistent with presence of Alzheimer's disease and this suggests that image based analysis of blood vessel structure might indicate the presence of Alzheimer's disease. Here we have deployed the standard measures of density and tortuosity for this analysis and have developed a new technique which is suited to analysis of the small branching structures to be found in these images. The new measure is formulated to have requisite invariant properties for this analysis.

In general, our initial result shows that the branching structure has become a major contributor to discriminate AD from a normal brain. In addition, a basic measure of tortuosity has shown some contribution to correct classification and has motivated us to increase the performance of the classification by finding a better way to analyse tortuosity.

Note that so far the study is in vitro and this study is sufficiently encouraging for translation to in vivo 3D MRI image analysis.

# References

[1] E. Boscolo, M. Folin, B. Nico, C. Grandi, D. Mangieri, V. Longo, *et al.*, Beta amyloid angiogenic activity in vitro and in vivo, *International Journal of Molecular Medicine*, 19:581, 2007.

[2] K. G. Mawuenyega, W. Sigurdson, V. Ovod, *et al.*, Decreased clearance of CNS β-amyloid in Alzheimer's disease, *Science*, **330**: 1774-1774, 2010.

[3] L. Wang, F. Beg, *et al.*, Large deformation diffeomorphism and momentum based hippocampal shape discrimination in dementia of the Alzheimer type, *IEEE Trans Med Imaging*, 26:462-70, 2007.

[4] P. A. Freeborough and N. C. Fox, MRI image texture analysis applied to the diagnosis and tracking of Alzheimer's disease, *IEEE Trans Med Imaging*,17: 475-9, 1998.

[5] G. Kojda and D. Harrison, "Interactions between NO and reactive oxygen species: pathophysiological importance in atherosclerosis, hypertension, diabetes and heart failure," *Cardiovascular research*, 43:652-671, 1999.

[6] K. E. Biron, D. L. Dickstein, R. Gopaul, *et al.*, Amyloid triggers extensive cerebral angiogenesis causing blood brain barrier permeability and hypervascularity in Alzheimer's disease, *PLoS One*, 6.8: e23789, 2011.

[7] Megalooikonomou, Vasilis, et al. "A representation and classification scheme for tree-like structures in medical images: analyzing the branching pattern of ductal trees in X-ray galactograms. *Medical Imaging, IEEE Transactions* on, **28(**4):487-493, 2009.

[8] Van Pelt, J., and R. W. H. Verwer. Topological properties of binary trees grown with order-dependent branching probabilities. *Bulletin of Mathematical Biology*, 48.2: 197-211, 1986.

[9] C. Y.-L. Cheung, Y. Zheng, W. Hsu, M. L. Lee, *et al.*, Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors, *Ophthalmology*, **118**: 812-818, 2011.

# Comparison of Threshold-Based Segmentation Methods on Pre- and Post-Therapy PET Scans

Michael Phillips[1]
Michael.Phillips.1@city.ac.uk
Sally F. Barrington[1]
Derek L.G. Hill[2]
Paul K. Marsden[1]

[1] Division of Imaging Sciences, PET Imaging Centre, King's College London, London, UK.
[2] IXICO Plc, Bioscience Innovation Centre, London, UK

**Abstract**

The aim of this study was to compare several threshold-based segmentation methods in delineating tumours on pre- and post- therapy PET scans both in terms of the volumes segmented and the effect on predicting survival. On a dataset of 14 patients with mesothelioma, different segmentation methods were found to correlate well with each other for pre- and post- therapy tumour volume (TV) and total lesion glycolysis (TLG) but did not correlated as well for absolute change and % change of TV and TLG between pre- and post- therapy images. This was also found for the effect of the TV and TLG on receiver operator characteristics (ROC) analysis on 6-month progression free survival (PFS) in where the segmentation method had little effect when assessing pre- or post- therapy TV and TLG but showed significant differences when using change of % change as a predictor of PFS.

# 1 Introduction

There are a number of different threshold-based region growing segmentation methods that have been used for segmenting tumours on PET scans. These range from using a fixed standardised uptake value (SUV) threshold of 2.5, found to distinguish between malignant and benign lesions [1], to methods which take into account the background in an image, such as those defined in the PERCIST criteria [2]. Segmentation can be used to delineate PET scans to obtain tumour volume (TV) and measures combining TV and SUVs, defined as total lesion glycolysis (TLG) [3]. Pre-therapy TV/TLG and the change in TV/TLG between pre- and post- therapy scans have shown promise as useful measures for predicting patient survival in mesothelioma and other cancers [4-8]. However, there is no designated, default method of segmentation and studies have used various threshold based methods such as a fixed 2.5 SUV threshold [4,5], fixed percentage of $SUV_{max}$ [6], PERCIST criteria method [7], and a method using $SUV_{mean}$ (named GRAB) [8]. The aim of this work is to compare these and other methods to see if the segmentation method makes a significant difference to the measurement of TV and TLG and, more importantly,

© 2014 . The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

whether it effects the prediction of response/survival in patients. This work has been completed using 14 pre- and post- therapy PET scans from a dataset investigating the use of Sorafenib as a therapy for mesothelioma [9].

# 2 Materials and Methods

## 2.1 PET Image Data

14 pre- and post- therapy PET scans of patients with mesothelioma were used for this study. All scans were acquired at the PET Imaging Centre in St Thomas' Hospital on either GE Discovery ST or GE Discovery VCT PET scanners (Waukesha, WI). Post-therapy scans were performed ~8 weeks after the start of treatment – a first line treatment of pemetrexed plus cisplatin chemotherapy, before a second line Sorafenib chemotherapy - and all patients had measurable disease, as defined by modified RECIST criteria [10]. PET image dimensions were 128 x 128 x 223, 267 or 311 (with voxel sizes of 5.47mm x 5.47mm x 3.27mm for all but one image with voxel sizes of 4.69mm x 4.69mm x 3.27mm). All PET images analysed were attenuation corrected using a smoothed CT dataset. Administered FDG dose ranged from 315 to 380MBq (median, 342MBq). The median time between administration of FDG and the start of the scan was 93min (range, 79-122min). Patients had a median age of 63 (range, 55-77) and 86% were male (12 male, 2 female). The study was approved by the UK National Research Ethics Service.

## 2.2 Disease Segmentation

All segmentation methods used 3-D region growing using 6-voxel connectivity in a software package called PETTRA (PET Therapy Response Assessor), a software tool designed using commercial software package MATLAB® 2012b (The MathWorks Inc., Natick, MA, 2000). Areas of segmentation were defined by an experienced consultant physician who decided which areas were disease and which were physiological uptake. In instances where the segmentation method clearly included areas of physiological uptake such as the liver, bladder or heart, restrictions were made using a given cubic area in which the segmentation must remain or which could remove unwanted physiological uptake, depending on which was easier. This was the same for all segmentation methods.

A fixed 2.5 SUV region growing segmentation method was used as the default method of segmentation due to its simplicity and objectivity. A fixed threshold based on a percentage of $SUV_{max}$ was also used with various percentages, as has been done in other studies [7]. The PERCIST method, which uses a formula of the mean of the background plus two standard deviations (S.D.), was used as a threshold with the background taken as a large volume of interest (VOI) over the centre of the liver. A large VOI was used to increase the robustness of the algorithm, rather than use a smaller VOI prone to more observer variability as proposed in the PERCIST guidelines [5]. This background measure was also used for other segmentations which included a background mean and/or S.D. Two methods using the $SUV_{mean}$ and background uptake were included, a method by Davis *et al*. (2006) which uses a relative percentage threshold of the $SUV_{max}$ – the background mean (this is then added to the $SUV_{mean}$ of the VOI) [11], and those defined by Nestle *et al*. (2005) and Nestle *et al*. (2007) which use a percentage of the $SUV_{mean}$ and a percentage of

the background added together [12,13]. Finally, the GRAB method, an adaptive method using the $SUV_{mean}$ and the background was used [14]. All the segmentation methods using the $SUV_{mean}$ as a starting point underwent an iterative process to continue recalculating the segmentation until the volume no longer changed or it had exceeded ten iterations with the starting $SUV_{mean}$ taken from a fixed 2.5 SUV threshold segmentation. Each segmentation method was used to segment disease on each of the 28 images (14 pre-therapy and 14 post-therapy). The TV and TLG was then calculated and compared for each segmentation method as was their resulting area under the curve (AUC) for receiver operating characteristics (ROC) analysis for 6-month progression free survival (PFS).

# 3 Results

## 3.1 Segmentations

Each pre- and post- therapy image was found to have some disease so all 28 images had a measure for TV and TLG. Segmentation using a threshold based on the percentage of the $SUV_{max}$ has been omitted from the results. A variety of fixed percentages were attempted but over the entire dataset they were found to either have too low thresholds at low percentages, which segmented almost the entire body, or too high thresholds at high percentages, which clearly segmented only a fraction of the disease in the image (Figure 1, (vi)). This was also true of some of the other segmentation methods tested, particularly when using different relative thresholds.

After removing segmentation methods which had problems segmenting many images, five region growing threshold segmentation methods were tested with thresholds based on: (i) a fixed 2.5 SUV [1], (ii) background uptake + 2 S.D., recommended by PERCIST [2], (iii) background mean + 0.10 ($SUV_{max}$ – background mean) where 0.10 is the chosen relative threshold factor, defined by Davis *et al*. (2006) [11], (iv) (0.15 * $SUV_{max}$) + background mean, defined by Nestle *et al*., (2005) [12], and (v) a threshold based on the adaptive GRAB method [14]. The threshold used in the GRAB method is given as:

$$\text{Threshold} = SUV_{mean} * \text{Threshold Factor}$$
$$\text{Threshold Factor} = 1 - ((SUV_{mean} - MNL) / (SUV_{mean} + MNL))$$
$$MNL = \text{Background Mean} + (\text{Background S.D.} * 3)$$

All background means and S.D.s were taken from a 201ml volume in the middle of the liver placed by an operator. This was true of all but one dataset which had disease present at the top of the liver so a smaller volume of 119ml was used instead to make sure no disease would alter the liver mean and S.D. When using the fixed 2.5 SUV threshold over the 28 images, 21 could be segmented with no restrictions. 4 pre-therapy and 3 post-therapy images had either the segmentation area restricted from physiological uptake or the physiological uptake removed from the segmented area. Physiological uptake included the heart, liver, spleen, bladder and bowel. All other segmentation methods were perceived to visually segment the diseased areas and had reasonable restrictions put in place to stop the segmentation of physiological uptake in comparison with the fixed 2.5 SUV method. A visual comparison of the segmentation methods on datasets 1 and 7 can be seen in Figure 1.

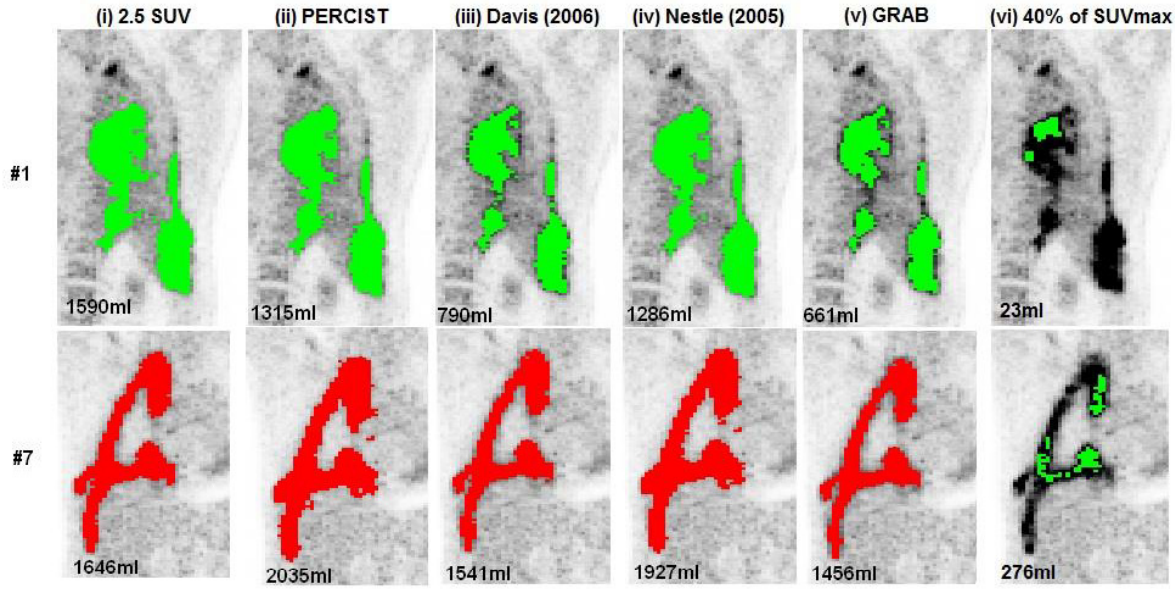



Figure 1: Coronal PET slices of segmentations of two datasets (1 and 7) by the five chosen methods: (i) fixed 2.5 SUV threshold, (ii) background mean + 2 background S.D., as recommended in the PERCIST guidelines, (iii) background mean + 0.10 ($SUV_{max}$ – background mean), defined by Davis *et al*. (2006), (iv) (0.15 * $SUV_{max}$) + background mean, defined by Nestle *et al*. (2005), and (v) the GRAB method. (vi) A threshold for 40% of the $SUV_{max}$ is also shown, as can be seen to the layman even on this 2-D coronal view, the method has clearly not segmented all of the disease.

## 3.2 Comparison of Segmentations

As can be seen in Figure 1, all segmentations cover the same areas of disease but differences in the threshold mean the TV can often be very different. A fixed 2.5 SUV threshold had a mean TV of 586ml over the 28 scans, while segmentations using PERCIST and the Nestle method had higher mean TV of 636ml. In comparison, the Davis method and GRAB generally produced lower TVs with means of 459ml and 327ml respectively. Compared to the fixed 2.5 SUV threshold segmentation method, the other four needed less restriction from or removal of physiological uptake in segmentation, however, it is worth noting that the PERCIST and GRAB methods needed to use more user initiated VOIs to do this over the dataset.

Table 1 shows the correlation between the fixed 2.5 SUV segmentation and the other four methods investigated for TV, TLG and the absolute change and absolute % change between them. The results show excellent correlation for all the segmentation methods for TV and TLG across all images but much weaker correlation when comparing the absolute change or % change between pre- and post- therapy images, with the PERCIST and GRAB methods having better correlation to using a fixed 2.5 SUV threshold than other methods.

| Segmentation Method | All Images | | Change | | % Change | |
|---|---|---|---|---|---|---|
| | TV | TLG | TV | TLG | TV | TLG |
| (ii) PERCIST | 0.946 | 0.984 | 0.677 | 0.962 | 0.943 | 0.974 |
| (iii) Davis *et al*. (2006) | 0.920 | 0.965 | 0.879 | 0.955 | 0.188 | 0.369 |
| (iv) Nestle *et al*. (2005) | 0.950 | 0.985 | 0.416 | 0.920 | 0.475 | 0.654 |
| (v) GRAB | 0.927 | 0.970 | 0.892 | 0.982 | 0.781 | 0.836 |

Table 1: Correlation between fixed 2.5 SUV threshold and other segmentation methods: shows the Pearson correlation coefficient (ppc) for TV and TLG over all 28 images and the absolute change and absolute % change over the 14 pre- and post- therapy images. For $p < 0.01$, $pcc > 0.479$ over all images and for $p < 0.01$, $pcc > 0.662$ for changes and % change.

### 3.3 Segmentation Effect on Predicting Response

A comparison of ROC AUC for 6-month PFS for each segmentation method shows that all segmentation methods produce similar results when using pre- and post-therapy TV and TLG to predict 6-month PFS with all methods producing AUC between 0.8 and 0.9 (with one exception, the Davis method (iii) with a pre-therapy TV AUC of 0.938), potentially caused by just one patient being categorised differently. For pre-therapy TLG all methods have identical ROC curves. However, when using change, or % change, between pre- and post- therapy TV or TLG the AUC ranges from 0.292 to 0.729. This suggests that while the segmentation method to obtain TV or TLG is unlikely to have a great effect when using pre- and post- therapy values to predict PFS, it is likely to have a greater impact when using the change, or % change, between pre- and post- therapy scans.

| Method | TV | | | | TLG | | | |
|---|---|---|---|---|---|---|---|---|
| | Pre | Post | Chg | % C | Pre | Post | Chg | % C |
| (i) 2.5 SUV | 0.875 | 0.812 | 0.396 | 0.438 | 0.875 | 0.833 | 0.313 | 0.396 |
| (ii) PERCIST | 0.875 | 0.875 | 0.646 | 0.437 | 0.875 | 0.854 | 0.500 | 0.396 |
| (iii) Davis | 0.938 | 0.833 | 0.396 | 0.417 | 0.875 | 0.833 | 0.313 | 0.375 |
| (iv) Nestle | 0.875 | 0.875 | 0.729 | 0.604 | 0.875 | 0.854 | 0.521 | 0.458 |
| (v) GRAB | 0.875 | 0.875 | 0.458 | 0.354 | 0.875 | 0.833 | 0.375 | 0.292 |

Table 2: Area under the ROC curve (AUC) for predicting progression free survival (PFS) at 6 months, using TV and TLG, for all segmentation methods. Chg = Change between pre- and post- therapy images. % C = % change between pre- and post- therapy images.

# 4 Conclusion

This study compares five threshold-based segmentation methods to delineate disease on pre- and post- therapy PET studies. On 14 patients with mesothelioma, all the other methods were shown to correlate well with the fixed 2.5 SUV threshold method and the segmentation method had little or no effect on ROC analysis when using TV or TLG to predict 6-month PFS. However, when using the change, or % change, between pre- and post- therapy TV and TLG there was a much lower correlation between segmentation methods and a greater impact on ROC analysis where the change, or % change, between TV and TLG produced vastly different AUC depending on the segmentation method.

# References

[1] A.C. Paulino and P.A. Johnstone. 2004. FDG-PET in radiotherapy treatment planning: Pandora's box? International Journal of Radiation Oncology Biology Physics, 59, 4-5.

[2] R.L. Wahl, H. Jacene, Y. Kasamon and M.A. Lodge. 2009. From RECIST to PERCIST: Evolving considerations for PET response criteria in solid tumours. Journal of Nuclear Medicine, 50 (1), 122S-50S.

[3] S.M. Larson, Y. Erdi, T. Akhurst et al. 1999. Tumor Treatment Response Based on Visual and Quantitative Changes in Global Tumor Glycolysis Using PET-FDG Imaging. Clinical Positron Imaging, 2, 159-171.

[4] P. Veit-Haibach, N.G. Schaefer, H.C. Steinert et al. 2010. Combined FDG-PET/CT in response evaluation of malignant pleural mesothelioma. Lung Cancer, 67, 311-317.

[5] M.K. Song, J.S. Chung, H.J. Shin et al. 2012. Prognostic value of metabolic tumor volume on PET/CT in primary gastrointestinal diffuse large B-cell lymphoma. Cancer Science, 103, 477-82.

[6] T.M. Kim, J.C. Paeng, I.K. Chun et al. 2013. Total lesion glycolysis in positron emission tomography is a better predictor of outcome than the International Prognostic Index for patients with diffuse large B-cell lymphoma. Cancer, 119 (6), 1195-202.

[7] H.Y. Lee, S.H. Hyun, K.S. Lee et al. 2010. Volume-Based Parameter of F-18-FDG PET/CT in Malignant Pleural Mesothelioma: Prediction of Therapeutic Response and Prognostic Implications. Annals of Surgical Oncology, 17, 2787-2794.

[8] R.J. Francis, M.J. Byrne, A.A. Van Der Schaaf et al. 2007. Early prediction of response to chemotherapy and survival in malignant pleural mesothelioma using a novel semiautomated 3-dimensional volume-based analysis of serial 18F-FDG PET scans. Journal of Nuclear Medicine, 48, 1449-58.

[9] S. Papa, S. Popat, R.Shah et al. 2013. Phase 2 Study of Sorafenib in Malignant Mesothelioma Previously Treated with Platinum-Containing Chemotherapy. Journal of Thoracic Oncology, 8(6), 783-87.

[10] M.J. Byrne and A.K. Nowak. 2004. Modified RECIST criteria for assessment of response in malignant pleural mesothelioma. Annals of Oncology, 15, 257-260.

[11] J.B. Davis, B. Reiner, M. Huser et al. 2006. Assessment of 18F-PET signals for automatic target volume definition in radiotherapy treatment planning. Radiotherapy & Oncology, 80, 43-50.

[12] U. Nestle, S. Kremp, A. Schaefer-Schuler et al. 2005. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. Journal of Nuclear Medicine, 46, 1342-8.

[13] U. Nestle, A. Schaefer-Schuler, S. Kremp et al. 2007. Target volume definition 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer. European Journal of Nuclear Medicine and Molecular Imaging, 34, 453-62.

[14] J.A. Boucek, R.J. Francis, C.G. Jones et al. 2008. Assessment of tumour response with (18)F-fluorodeoxyglucose positron emission tomography using three-dimensional measures compared to SUVmax – a phantom study. Physics in Medicine and Biology, 53, 4213-30.

# Speeding Up Active Mesh Segmentation by Local Termination of Nodes

Carl J. Nelson[1]
carl.nelson@dur.ac.uk
Martin Dixon[2]
martin.dixon@durham.ac.uk
P. Philippe Laissue[3]
plaissue@essex.ac.uk
Boguslaw Obara*[1]
boguslaw.obara@dur.ac.uk

[1] School of Engineering & Computer Science
Durham University
[2] School of Biological and Biomedical Sciences
Durham University
[3] School of Biological Sciences
University of Essex

**Abstract**

This article outlines a procedure for speeding up segmentation of images using active mesh systems. Active meshes and other deformable models are very popular in image segmentation due to their ability to capture weak or missing boundary information; however, where strong edges exist, computations are still done after mesh nodes have settled on the boundary. This can lead to extra computational time whilst the system continues to deform completed regions of the mesh. We propose a local termination procedure, reducing these unnecessary computations and speeding up segmentation time with minimal loss of quality.

## 1 Introduction

Deformable models [6] are a group of algorithms used to segment images, particularly images with weak or missing boundaries. Such deformable models, also known as active models and snakes, are used heavily in biomedical imaging segmentation where low contrast, high noise or obscured edges make other segmentation techniques unsuitable. Deformable models come in several forms and have been used for multiple modalities and scenarios (see [5] for a comprehensive review). Here we introduce and briefly describe the field limited to active meshes, three-dimensional polygonal-faced meshes that can be deformed in a discrete manner to segment objects of interest [7].

### 1.1 Active Meshes: a Quick Background

Active meshes are a subset of deformable models that commonly use a triangular-faced mesh to segment an object of interest from an image with weak or missing boundaries (see [4] for an early example). The use of a mesh with distinct nodes, or vertices, and a known connectivity, through face or edge information, allows computationally simple and speedy calculations to be performed over the whole mesh. To find the segmentation shape the mesh is deformed in one of several ways. The use of physics-based forces allows a new position to be calculated at each iteration by Euler-Lagrange mechanics [8]; however, there have been concerns

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

that optimality of such systems is not guaranteed [1]. A more common system is to provide each vertex with an energy and, for each iteration, identify some trial points for the vertex to move to; one of these points is then accepted by some optimisation method, the simplest of which is a greedy algorithm [10]. Many of these systems require a balloon force [3] or another force (e.g. [11]) to attract the shape towards the edges from any initialisation. As with many techniques each variant has its advantages and drawbacks and the optimal choice for a particular case is dependent on the experiment and researcher.

### 1.2 Termination of Meshes

Deformable models tend to work on an iterative system with each iteration deforming the model to minimise energies or balance forces, eventually the mesh reaches an optimal state and the desired object is segmented. With many optimisation techniques there may be no definite and/or optimal termination; as such, many systems are provided with an explicit stopping criterion in order to terminate the process. One key feature of termination criteria that are mentioned in the literature is that they focus on the overall mesh, i.e. they terminate in a global fashion, e.g. the percentage of 'inactive' vertices [8] or a maximum number of iterations [9]. We have been unable to find any articles in the literature that use a system of local vertex termination, that is to say, each vertex is deformed as part of the mesh until it reaches some independent termination criteria. In this article we propose a heuristic system for local vertex termination during deformation. We show that this method of termination can speed up the overall process of segmentation by testing the algorithm on synthetic data and also demonstrating the system on real 3D images.

## 2 Methods

For the purpose of this article we have used a simple balloon-driven, energy minimisation mesh (based on [2]). The mesh is initiated as an icosahedron of user-defined size. For each iteration, every vertex is assigned a new 'more optimal' position selected, by a simple greedy algorithm, from a 3D neighbourhood. A steepest gradient descent method is employed to link the image gradient to the size of the steps available. After each iteration of deformation the mesh undergoes local resampling, again based on [2]. During each iteration, all vertices are checked to see if they should be terminated (described fully below). Deformation continues until a termination criterion is met; this criterion will be described for each experiment in Results. All codes were written in MATLAB 2012b (The MathWorks, Inc., US) on a Windows 7 64-bit PC running an Intel Core i5-2320 CPU (3.0 GHz) with 6GB RAM.

### 2.1 Local Termination of Vertices

In order to terminate vertices we identify, at the end of each deformation, a list of vertices to be considered 'terminated'. During the next iteration we only deform vertices that are not terminated, decreasing computation time per iteration. Our system checks whether or not a vertex should be considered terminated after every iteration; this prevents a vertex being turned off due to, for example, a large noise signal, and being prevented from re-initiation as the mesh around it deforms. We define whether or not a vertex should be terminated using the following cost function,

$$C_v = \alpha C_I + \beta C_G + \gamma C_T, \tag{1}$$

where $\alpha, \beta$ and $\gamma$ are user-defined parameters such that $\alpha + \beta + \gamma = 1$. The first two terms, $C_I$ and $C_G$, are the value of the normalised image and gradient magnitude image,

respectively, calculated through trilinear interpolation. Finally, $C_T$ is one if the vertex was terminated during the last iteration and zero otherwise. We note that more advanced cost functions could look at the history or neighbourhood of the vertex in more detail but we have, so far, only looked at basic, faster cost functions. In order to make this cost-based method more robust we use a combination of the cost at the vertex and the average cost of its direct neighbourhood. So $T_v$, the choice to terminate a vertex or not, is given by,

$$T_v = \begin{cases} \texttt{TRUE} & \text{if } \mu C_v + (1-\mu)\frac{\sum_j C_j}{N_v} > c, \\ \texttt{FALSE} & \text{otherwise;} \end{cases} \tag{2}$$

here, $N_v$ is the size of the neighbourhood around the vertex, $v$, $C_j$ is the cost at a neighbour, $j$, and $\mu$ is a user-defined parameter. The user-defined cost-level, $c$, is a threshold above which a point is terminated.

## 2.2 Measure of Quality

To determine segmentation 'quality', $Q$, of synthetic images we have used a simple metric comparing the segmented mesh to the binary 'ground truth' image, $G$. The resultant mesh is voxelised, $B$, and the measure of quality is calculated as $Q = B \cap G / B \cup G$. If the mesh extends outside of the ground truth the denominator increases and if the mesh has not fully extended to the ground truth boundaries then the numerator decreases. An ideal segmentation will always have a quality of one. Any segmentation that over or under segments will have a quality less than one; over segmentation and under segmentation in different regions cannot balance out using this metric.

# 3 Results

In order to clearly show when and where local termination may be effective for active mesh segmentation we have run experiments on five noise-free synthetic shapes. We also compare the results of globally terminated and locally terminated segmentation on real data of dehisced mature pollen grains.

## 3.1 Synthetic Data

We have chosen five synthetic shapes to represent various scenarios found in biological and biomedical images: a sphere, a cube, a hollow, rectangular cuboid, a tube and a weeble (a cone placed on top a hemisphere). We have chosen to demonstrate on noise-free data to show the differing effect of local termination on different shape scenarios, where we feel the difference is most significant. User-chosen parameters for local termination were used for each shape. The origin for segmentation was user-chosen based on manual selection of an approximate centroid. For all the control segmentations, i.e. segmentations using global termination, we terminated the procedure based upon the Haussdorf distance before and after each iteration. For each step the Haussdorf distance between the set of vertices before deformation and resampling and after were compared. If the Haussdorf distance was below a user-defined value then the shape was considered to be terminated.

Table 1 shows the time taken to segment and final segmentation quality for these shapes under global and local termination procedures. As can be seen there is a dramatic decrease in segmentation time with only a moderate loss of quality. We note that shapes that are fairly round or with a surface fairly evenly distributed from the point of initiation, i.e. the weeble and sphere, reap the smallest benefit in terms of time. The very slight increase in

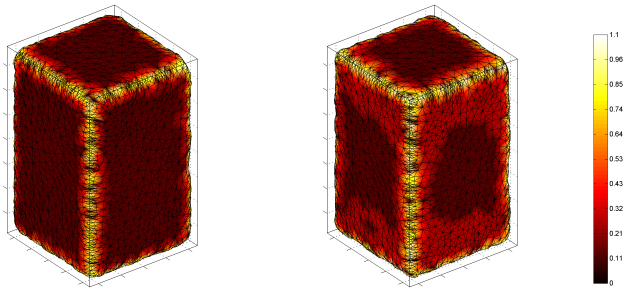
| Shape (Vertices) | Global Termination | | Local Termination | | Reduction | |
|---|---|---|---|---|---|---|
| | Time(s) | Quality | Time(s) | Quality | Mean Time | Quality |
| Sphere (4,342) | 186.1(0.1) | 0.9564 | 70.0(0.4) | 0.9595 | 62.4% | -0.3% |
| Cube (4,281) | 336.0(0.7) | 0.9241 | 69.8(0.2) | 0.8730 | 79.2% | 5.5% |
| Weeble (6,606) | 251.4(0.6) | 0.9589 | 118.2(0.5) | 0.9647 | 53.0% | -0.6% |
| Tube (6,617) | 265.5(0.7) | 0.9580 | 40.2(0.07) | 0.9287 | 84.9% | 3.1% |
| Cuboid (3,694) | 124.8(0.3) | 0.9142 | 59.3(0.03) | 0.8738 | 52.5% | 4.4% |

Table 1: Table showing reduction in time and quality of segmentation using local termination procedures. Times are taken from an arithmetic mean of ten full segmentations; the number in standard brackets is the standard error upon the mean to one significant figure. Number of vertices, taken from global termination results, are shown in brackets after shape names.

quality seen in some shapes is caused by the 'wiggle' of vertices across the surfaces due to the multidirectional nature of our method: as the Haussdorf distance measures the minimum distance between one vertex before deformation and its closest match after, it is conceivable that, as the mesh settles on the surface of the shape, vertices 'wiggle' in a Brownian motion-like manner across the surface, leading to resampling of particular regions of the surface, hence keeping the Haussdorf distance high. We should also note that we tracked the quality of the segmentation for each iteration (data not shown) and found that the maximum quality was not at the point of termination but a few iterations before termination due to the same reason. Figure 1 shows distance maps from the segmentation mesh of the hollow cuboid. The colouring at that area of the surface shows the average distance between that region of the mesh and the true boundary of the synthetic shape.

## 3.2 Noisy Data

In order to test our local termination data on images more similar to those found in biological and biomedical modalities we applied pseudo-random noise to our hollow cuboid image. Under these noisy conditions the local termination method lead to a decrease of 50 ($\pm$10)% in run time with only a 7.9 ($\pm$0.5)% decrease in quality compared to the global termination method. The average peak signal to noise ratio for the global termination trials (n=10) was 73.1685 ($\pm$0.0002) and for the local termination trials (n=10) was 73.1688 ($\pm$0.0003); peak signal to noise ratio was calculated as $\texttt{maximum}(I_s)/\texttt{variance}(N)$, where $I_s$ is the noise-free image and $N$ the noise applied to the image.




(a) (b)

Figure 1: Local Termination of Vertices Reduces Segmentation Time with a Small Reduction in Quality. (a) Distance difference map of segmentation using global termination. The segmented mesh is compared to a ground truth image and the colours represent the distance between the mesh and ground truth. (b) Distance difference map of local termination-based segmentation. The scale bar applies to both subfigures. Units are in voxels.

| Local Termination... | Time(s) | Quality | Reduction from Global | |
|---|---|---|---|---|
| | | | Mean Time | Quality |
| with Random Benefit | 54(2) | 0.9737(0.008) | 79.4% | -1.6% |
| with Double Take | 49.54(0.08) | 0.9733 | 81.3% | -1.6% |

Table 2: Table showing the results of variations upon the basic cost-based local termination method. Note that the Quality of Local Termination with Random Benefit segmentation is stated as a mean (standard error) of ten repeats due to the use of a random value.

### 3.3 Improving Segmentation Quality

As shown above, local termination of active mesh segmentation can lead to a dramatic reduction in computing time; however, segmentation within biological and biomedical imaging requires as high a precision as possible. Here we show two slight modifications of our local termination scheme that help to improve the quality of segmentation. The results of these methods are shown in Table 2 for the synthetic tube; we have chosen this shape as there was the largest effect on reduction in time and a small decrease in quality. Tube-like structures are also heavily prevalent in biomedical images.

The first variation subtracts a pseudo-random value between 0 and 0.1 from the termination cost before comparison to $c$; this causes some vertices that should be terminated this iteration to be left on for at least one more iteration. As this is random it should lead to a slightly improved segmentation quality at the end result. With this method there is, however, a chance that no iteration will have all vertices terminated and, as such, a global termination criterion must be employed; for consistency we have chosen the Haussdorf distance as used above. In Table 2, this method is referred to as 'with Random Benefit'. The second method ignores the first termination of a vertex, i.e. a vertex must be classified as terminated for two iterations before it is actually terminated. The final termination criterion for this method is that when all vertices are terminated the segmentation is considered completed. In Table 2, this method is referred to as 'with Double Take'.

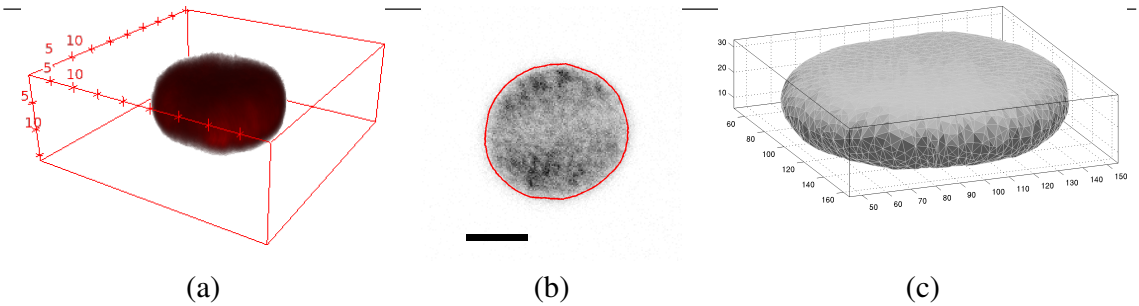
Table 2 shows that both variations keep the reduction in segmentation time low but improve the segmentation quality; in fact, both methods improve the quality beyond the original segmentation system. It is likely that this difference in quality is due to the particular stopping criterion for global termination used throughout this paper. As noted above the maximum quality for segmentation is often not the final quality when using the Haussdorf distance.

### 3.4 Biological Data

The proposed approach has been tested and validated on 3D CLSM images of DAPI-stained dehisced mature pollen grains obtained from wildtype *Arabidopsis thaliana* plants of the Columbia (Col-0) ecotype (Figure 2). On average, the proposed local termination approach has been able to speed up the overall process of pollen grains segmentation by 51%.

## 4 Conclusion

We have demonstrated the effect of local termination of vertices on the segmentation of shapes using active mesh systems. We have shown how the segmentation of long thin shapes is sped up dramatically by terminating those vertices that are 'complete' and no longer deforming the mesh in that region. We have chosen to show a basic cost function with some interesting improvements upon the basic idea; however, we are currently looking into more advanced ways of deciding whether or not a vertex is completed and, in general, the idea



(a) (b) (c)

Figure 2: Mature pollen grain: (a) 3D volume rendering. (b) 2D contour and (c) 3D mesh representation of the segmentation results. Scale bar corresponds to $10\mu m$.

of a cost function can be easily extrapolated for many imaging scenarios. Scenarios where this system will have the greatest effect are the segmentation of long, tubular structures, e.g. blood vessels, nerves or pathways such as the branchial network of the lungs, and by combining local termination with a topologically adaptive active mesh we could segment complicated tubular systems in a fraction of the time of the traditional methods. We are currently expanding application of the algorithm on biological and medical datasets.

# References

[1] AA Amini, TE Weymouth, and RC Jain. Using dynamic programming for solving variational problems in vision. *Pattern Anal.*, 12(9):855–67, 1990.

[2] AJ Bulpitt and ND Efford. An efficient 3D deformable model with a self-optimising mesh. *Image Vis. Comput.*, 14:573–80, 1996.

[3] LD Cohen. On active contour models and balloons. *CVGIP*, 53(2):211–8, 1991.

[4] H Delingette. Adaptive and deformable models based on simplex meshes. *Proc. 1994 IEEE Work. Motion Non-rigid Articul. Objects*, pages 152–7, 1994.

[5] L He, Z Peng, B Everding, X Wang, CY Han, KL Weiss, and WG Wee. A comparative study of deformable contour methods on medical image segmentation. *Image Vis. Comput.*, 26(2):141–63, 2008.

[6] M Kass, A Witkin, and D Terzopoulos. Snakes: Active contour models. *Int. J. Comput. Vis.*, 331:321–31, 1988.

[7] D Molloy and PF Whelan. Active-meshes. *Pattern Recognit. Lett.*, 21:1071–80, 2000.

[8] J-Y Park, T McInerney, D Terzopoulos, and M-H Kim. A non-self-intersecting adaptive deformable surface for complex boundary extraction from volumetric images. *Comput. Graph.*, 25(3):421–40, 2001.

[9] X Wang, L He, and WG Wee. Deformable Contour Method: A Constrained Optimization Approach. *Int. J. Comput. Vis.*, 59(1):87–108, 2004.

[10] DJ Williams and M Shah. A fast algorithm for active contours and curvature estimation. *CVGIP*, 1992.

[11] C Xu and JL Prince. Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.*, 7(3):359–69, 1998.

# Segmentation of Individual Cells in Phase Contrast Microscopy Images

Jindřich Soukup[1,2,3]
soukupj@utia.cas.cz

Michal Lašan[2]

Filip Šroubek[1]

[1] Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic
[2] Faculty of Mathematics and Physics, Charles University in Prague
[3] Institute of complex systems, FFPS, University of South Bohemia in Ceske Budejovice

**Abstract**

Phase contrast microscopy is widely used for live cell imaging. Automatic computer processing of phase-contrast images is complicated, particularly because of the halos that are present around the cells. We introduce a new algorithm able to segment individual cells by using the halos. The halos between cells are detected by thresholding and skeletonization, and gaps in the edges are filled in using Dijkstra's algorithm. The result of the algorithm was compared with the results of manual segmentation performed by a human expert and evaluated on level of precision of individual cells' edges. The source code and test data, including the manual segmentation are made available to the public.

## 1 Introduction

Nowadays, computer processing of microscopic images is becoming more common as a part of research in the field of cell science, replacing slow and time-consuming manual processing by human experts. A challanging task is for example to process images of live mammalian cancer cells from a phase contrast microscope. The cells have irregular shape, appear semi-transparent in the images (their inner structure is visible), and the colour of their inner parts is quite similar to the background colour. A characteristic feature of this microscopy method are light areas around the cells, so called halos (see Fig. 1). All this complicates the automatic computer processing of such images.

When processing the time-lapse images of live cells, the main goal is to describe the movement and behaviour of individual cells. From the image processing point of view, this means segmenting individual cells.

We distinguish two categories of segmentation methods that process phase contrast microscopy images: Methods [1,2] that separate background from contiguous cells, and methods [3,4,5,6] that segment individual cells from colonies (densely packed cells touching each other) or even from cell monolayers (cells create a contiguous layer) which is much more difficult.

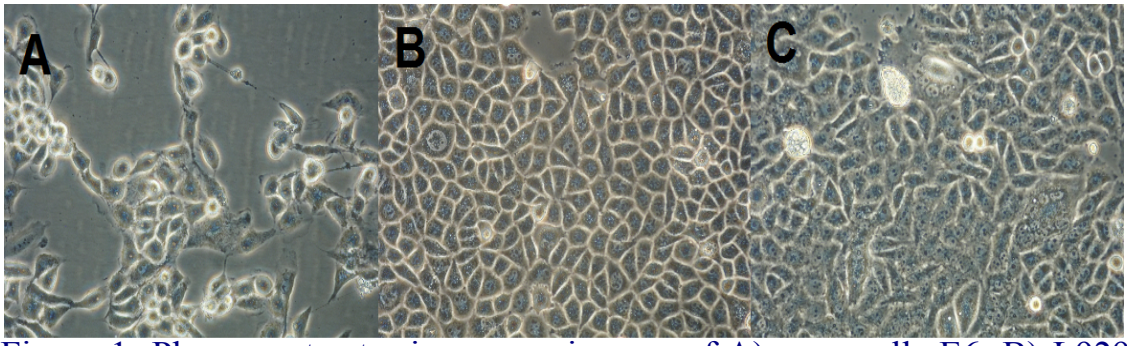



Figure 1: Phase contrast microscopy images of A) vero cells E6, B) L929 mouse fibroblast, C) HeLa cerivix epitheloid carcino cells.

Methods [3,4] were applied to contiguous cells, methods [5,6] were tested only on images containing separated cells with small spaces in between.

All the above-mentioned methods have their advantages, as well as disadvantages (require labeled database, high computational demands, lengthy parameter tuning, ...) and may fail when processing certain types of data (noise, blurring, presence of impurities).

Our work offers an alternative way to segment individual cells. It is based on the presence of halos around individual cells and it can process images of many contiguous cells or images of cell monolayers. To evaluate the algorithm's performance we developed a numerical evaluation on the level of individual cells based on precision, recall and F1 measures.

To provide a baseline for testing performance of previously mentioned methods or any future methods we provide the method's source code, test data and manual segmentation by human experts on our website http://zoi.utia.cas.cz/IndividualCells.

Together with the algorithm itself we have developed a software tool for editing the segmentation results (merging or splitting the cell regions), so that any mistakes of the algorithm can be easily and quickly corrected, making the algorithm more usable.

The rest of the paper is organized as follows: Sec. Methods describes our algorithm, and Sec. Results shows the precision of our algorithm when compared to human experts.

# 2 Methods

The goal of our method is to find individual cells in the images, or more precisely their borders. The algorithm consists of several stages. In the first stage, an incomplete set of edges between cells is created by thresholding and subsequent skeletonization. The aim of the second stage is to fill in the missing parts of the edges by connecting the loose ends of the skeleton. The third stage adjusts the resulting edge map by merging too small areas to the neighbouring cell areas. In the following text we first explain the main

principles of the method and then describe modifications that further improve the method.

The main characteristic of phase contrast microscopy images is that cells are often surrounded by a white area – the halo see Fig. 2A. Halos are also present in areas where two cells touch each other. This is used by the first stage of our algorithm. First we slightly blur the image (by gaussian with standart deviation 10) and divide it into lighter and darker areas (see Fig. 2B) using Otsu's automatic thresholding method [7]. The blurring ensures that the thresholding does not select the occasional light points inside the cells. The lighter area therefore consists mostly of the halos. The result of the thresholding is subsequently subjected to skeletonization (see Fig. 2C). We use a modified version of skeletonization [8], which reduces the number of short branches of the main skeleton.

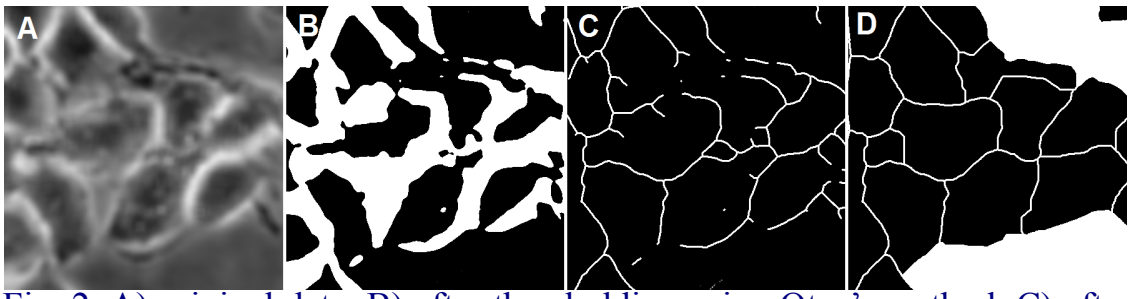



Fig. 2: A) original data, B) after thresholding using Otsu's method, C) after skeletonization, D) result of the method - white region is background, white lines are found cell borders.

We presume that most of the loose ends of the skeletonization results from the halo being less distinctive in the given area, and so in the second stage of the algorithm we connect these loose ends to other skeleton fragments. In order to make these connections closer to reality, we try to make them as short as possible and keep them in the brightest areas. For this we use Dijkstra's algorithm for finding the shortest path in a graph [9]. We represent every pixel as a vertex $V$ of a graph and edges between the vertices connect neighbouring pixels using 8-neighbourhood connectivity. The edge weight is calculated based on the pixel intensity $I$ as $f( I(V_1)+I(V_2) )$, with $f$ being a function that assigns lower value to the lighter pixels, e.g. $f(x) = -x$ or $f(x) = -log(x+1)$. For our experiments we used the latter one. For the diagonal edges we multiply the weight by $sqrt(2)$. Dijkstra's algorithm runs from each loose end and finds the nearest part of the skeleton. At the end of this phase we have an outline of individual cell edges.

At the end we check whether the results correspond with our knowledge of the studied cells. For this we use several criteria which we do not discuss due to lack of space but the main is the cell size. If the final border segmentation shows a smaller area than this criterion allows, we fill in this area completely and skeletonize it once more. In most of the cases the smallest possible size of a cell corresponds to the size of a cell in the mitotic phase (cells shrink during this phase). Final result can be seen at Fig. 2D.

The algorithm can be substantially improved when prior to connecting the loose ends of skeleton we first label the image pixels as belonging to cells or background. There are several algorithms available for this task– we use the SegTool programme [2] which is fast and quite accurate. This modification prevents drawing false edges between clusters of cells that do not touch each other. It also makes searching for optimal connections in Dijkstra's algorithm faster.

Other modifications adjust the process of searching for optimal connections using Dijkstra's algorithm. The algorithm has a natural tendency to use connections displayed in Fig 3A (the red or yellow connections). To achieve the green connection, we use two ad hoc rules. The first one states that the connecting line must leave the contour. By the conture we mean the result of the Otsu thresholding as seen at Fig. 2B. This rule prevents the red connection, see Fig. 3A. The second rule defines the direction in which a connection line is searched, see Fig. 3B.

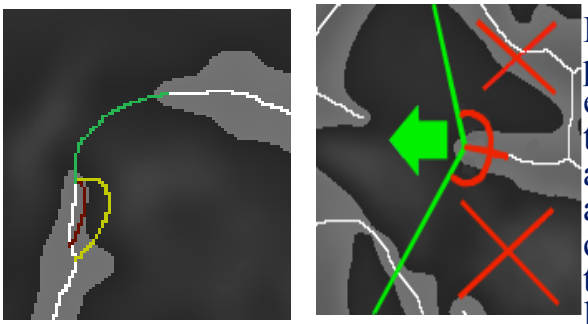


Figure 3: A) Illustration of the problems with connecting skeleton ends. The green connection line is the preferred one. Without additional modifications the algorithm creates the red or yellow connection line. B) Illustration of the rule for direction of connection line search.

The algorithm has several inner parameters. These are ad hoc rules that enhance the performance of the algorithm, such as the amount of initial image blurring and two rules (in Sec. Alg. Modifications) for searching connections between contour ends. These parameters need to be tuned only once to achieve the most successful segmentation and they are fixed in subsequent runs of the algorithm. The minimal size of cells depends on the type of scanned cells and is defined by the user.

# 3 Results

When comparing results of individual cell segmentation with ground truth (in our case manual segmentation done by human experts), we can distinguish three types of errors in the segmentation results: A) misaligned cell borders, B) two or more cells merged into one region C) one cell split into two or more regions. Boundaries between these three error categories are not well defined, combination of merging and splitting could sometimes results in cell border misalignment and vice versa. Therefore our evaluation method measures these three types of errors together.

For each cell from the manual segmentation we found an area from our algorithm that had the largest overlap. For all these pairs we measured the size of the overlapping area – true positive (TP), the size of the area in the

manual segmentation apart from the overlap – false negative (FN), and the size of the area in our segmentation apart from the overlap – false positive (FP). We added up the sizes of TP, FP and FN of all the cells pairs and converted them to precision (P = TP/(TP+FP) ), recall (R = TP/(TP+FN) ) a F1 (F1 = 2PR/(P+R)) measures.

The algorithm was implemented in Java and Matlab. The CPU time (using Dual core 2.3 Ghz) for segmentation of 4 MPixel image was about 30 seconds. The algorithm was tested using three different types of cells: HeLa cerivix epitheloid carcino cells (4 images), L929 mouse fibroblast (1 image), and vero cells E6 (3 images). The images were taken using Olympus X51S8F-3 microscope with magnification 20x and resolution 2288x1712. In total 8 images were considered on which were almost 2000 cells. We compared the algorithm results with results of manual segmentation done by an expert. The mean precision values for all of the test images are P = 0.65, R = 0.73, and F1 = 0.68 (cell-background labelling was provided by SegmentationTool) and P = 0.66, R = 0.74, and F1 = 0.69 (cell-background labelling was taken from manual one by merging all the cells). Number of the cells that our algorithm found is about 13% higher than real number.

As we have already mentioned, the results of the automatic segmentation algorithm can be improved (connecting or dividing areas and shifting edges) by our software tool. The quality of segmentation therefore influences only the amount of work needed to correct the mistakes.

The obtained results are graphically represented in Fig 3. Yellow signifies under-segmented areas (where our algorithm connected two real cells into a larger one), light blue is for over-segmented areas (containing extra borders). Green signifies areas that were segmented accurately and dark blue is used for background without any cells on it.

For comparison, we tested ordinary watershed method on our dataset. Precision of this method is much worse (P = 0.87, R = 0.26, F1 = 0.40) mainly due to oversegmentation (algorithm found 22-times more cells than present in the image).

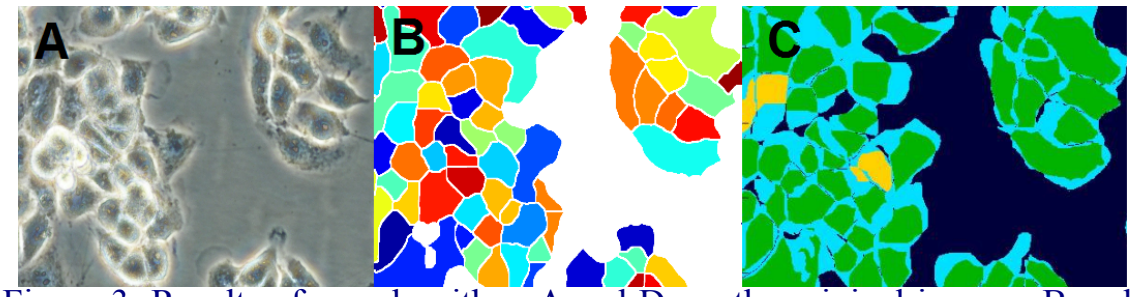



Figure 3: Results of our algorithm. A and D are the original images. B and E are examples of segmentation by our algorithm. C and F show correspondence with manual segmentation (green areas are correct, light blue ones are under-segmented and yellow are over-segmented. Dark blue is used for background).

# 4 Discussion and conclusions

We introduced a new method for segmentation of individual cells in images acquired by phase contrast microscopes. Unlike previous segmentation methods, it uses the presence of halos around individual cells.

One of the weaknesses of this algorithm is that it processes the individual images from time series separately. The algorithm could be improved by connecting tracking and segmentation. Further improvement could be achieved by using methods, such as active contours or level-sets, since the results of our algorithm could provide a good initialization for them.

Our method is currently being tested at Working place of tissue culture - certified laboratory, which researches on new potential materials for implants in vitro.

# Acknowledgment

The results of the project LO1205 were obtained with a financial support from the MEYS under the NPU I program, CENAKVA CZ.1.05/2.1.00/ 01.0024. The authors acknowledge the support of the GA JU 134/2013/Z, GAUK, grant No. 914813/2013 and grant GAČR No. 13-29225S.

# References

[1] Ersoy, I. et al..: Cell segmentation using Hessian-based detection and contour evolution with directional derivatives, *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on,* **2008**, *2008*, 1804-1807

[2] Soukup, J. et al.: Segmentation of Time-Lapse Images with focus on Microscopy Images of Cells, *17th International Conference on Image Analysys and Processing - ICIAP 2013, Springer-Verlag,* 2013, *8157, Part II*, 71-80

[3] Wang, X. et al.: Cell segmentation for division rate estimation in computerized video time-lapse microscopy, *Proceedings of SPIE, Spie,* **2007**, *6431*, 643109-643109-8

[4] Pan, J. et al.: Heterogeneous Conditional Random Field: Realizing joint detection and segmentation of cell regions in microscopic images, *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE,* **2010**, 2940-2947

[5] Li, K. et al.: Cell population tracking and lineage construction with spatiotemporal context, *Medical image analysis,* **2008**, *12*, 546-66

[6] Yin, Z. & Kanade, T.: Restoring artifact-free microscopy image sequences, *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on,* **2011**, 909-913

[7] Otsu, N.: A threshold selection method from gray-level histograms, *Automatica,* **1975**, *20*, 62-66

[8] Hilditch, C.: Linear Skeleton from square cupboards, *Machine intelligence,* **1969**, *6*, 403-409

[9] Dijkstra, E.: A note on two problems in connexion with graphs, *Numerische Mathematik, Springer-Verlag,* **1959**, *1*, 269-271

# Utilizing phase retardation features for segmenting cells in phase contrast microscopy images

Thomas A. Nketia
thomas.nketia@eng.ox.ac.uk
Jens Rittscher
jens.rittscher@eng.ox.ac.uk
Alison J. Noble
alison.noble@eng.ox.ac.uk

Department of Engineering Science
University of Oxford
Parks Road
Oxford
OX1 3PJ

**Abstract**

Label free imaging, especially, phase contrast microscopy, plays an important role in high-content microscopy, in particular live cell screening. We aim to develop a fast and robust segmentation algorithm that enables the extraction of accurate morphometric measurements at a single cell level. Here we make use of a recently introduced image restoration method, which captures intrinsic features of phase contrast images. The resulting features are aggregated into superpixels and then grouped into objects using standard clustering methods.

The resulting method is computationally effective and only requires a minimal amount of user annotation. Proposed method is tested on a set of 10 phase contrast images of cervical cancer cell colonies of the HeLa cells. In order to compare our results with the literature, we report a Tanimoto coefficient of 0.94 for the set of images that were used in this study. The reported results indicate that our method does compare with the current state of the art.

## 1 Introduction

High-content microscopy plays an increasingly important role in drug development and basic life science research. Here we are particularly interested in enabling live cell imaging [6] experiments that reveal more detailed information about cellular function than standard end point assays. Given the volume and complexity of these image data sets there is a need for analysis tools that measure parameters such as overall cell morphology, measure parameters associated with cell division, motility, mechanisms of death and population statistics. Label free imaging, in particular, phase contrast microscopy, plays an important role since it eliminates the need for potentially cytotoxic fluorescent markers.

One fundamental task in quantitative biological imaging is cell segmentation. While a number of sophisticated and robust segmentation techniques have been developed for fluorescent microscopy, cell segmentation in phase contrast image sequences is still challenging. Artefacts that are introduced by the specific optics, observed as halo and shade off effects, as

© 2014 . The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

well as the lack of clear boundaries among cluttered cells, poor contrast between cells boundaries need to be addressed. Ali *et al.* [2] uses local phase analysis and level set techniques to automatically extract cells from background. Wang *et al.* [8] apply machine learning to learn the specific edge profiles observed in phase contrast data whereas Arteta *et al.*[3] employs a structured learning scheme with dot annotations over non-overlapping extremal region features. We are particularly interested in a recently developed image restoration method [7, 9] that produces phase retardation features, which are intrinsic to phase contrast microscopy.

Here we are presenting a weakly supervised method that makes direct use of the phase retardation features introduced in [7]. After extracting the phase retardation features [9], the pixel features are aggregated into SLIC superpixels [1]. Subsequently, we utilize standard clustering and watershed segmentation to find an accurate delineation of the cell boundaries. Weak *a priori* knowledge about the location of the individual cells is being used to achieve robustness of the overall method. Details of our approach are presented in Section 2.

Results from our experimental evaluation are presented in Section 3. Our experiments show that the proposed method gives a good segmentation of cells, achieving an average Tanimoto coefficient (TC) of 0.94. Common failure modes, as for example the risk of over-segmentation are discussed in more detail. In the conclusion section, Section 4 we will outline why the proposed method is well suited for tracking cells from frame-to-frame in a time lapse sequence.

# 2 Method

Extracting accurate morphometric information for each individual cell requires reliable detection of object boundaries. Given the nature of phase contrast microscopy, the appearance of such object boundaries depends on the specimen itself as well as the local imaging conditions. This motivates why we utilize the phase retardation features, which are then aggregated into superpixels. Clustering is used to label the superpixels, which results in an over-segmentation of the image. The identification of the individual cell bodies is explained in Section 2.3.

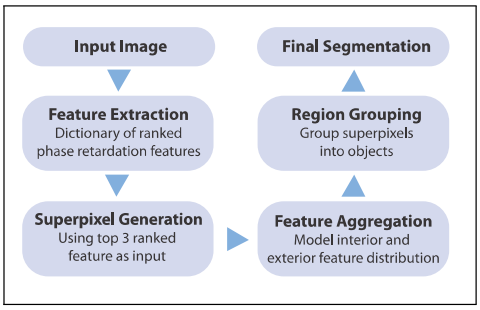
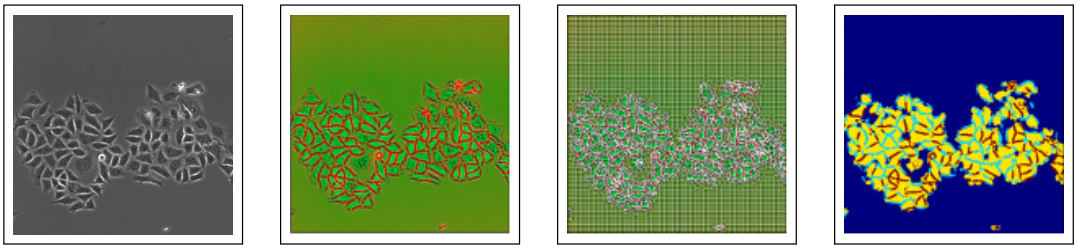



Figure 1: **Algorithm Overview.**

## 2.1 Feature extraction

Phase contrast microscopy is essentially a contrast-enhancing technique that employs an optical mechanism to translate small variations in phase (diffraction pattern) into corresponding changes in amplitude, which can then be visualized as differences in image intensity. This image formation process can therefore be exploited in an image reconstruction method to obtain features for each pixel based on the diffraction and hence the phase retardation experienced by each pixel element. We adopt the feature extraction method outlined in [7, 9] which proposes a dictionary-based imaging model for an $N$-pixel grayscale image as:

$$g = \sum_{m=0}^{M-1} H_m \Psi_m, \quad s.t.\ \Psi_m \geq 0, \tag{1}$$



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

Figure 2: Initial processes resulting in cluster regions: (a) Original phase contrast image; (b) The top three(3) principal features for each pixel mapped onto RGB of a color image; (c) SLIC superpixel over-segmentation; (d) Result of GMM clustering

where $g$ is the vectorized image intensities ($N \times 1$) expressed as a linear combination of a dictionary of $M$ bases $\{H_m\}_{m=0}^{M-1}$. The vectorized coefficients, $\Psi_m$ for each pixel corresponds to a diffraction pattern with phase retardation $\theta_m = \frac{2\pi m}{M}$. Hence, the $n^{th}$ row of an $H_m$ ($N \times N$) matrix corresponds to the vectorized kernel function obtained from a point spread function around pixel location $x_n$. The vectorized coefficients for each grayscale pixel are then utilized as the $M$-dimensional feature vector for that pixel.

The principal features are obtained from the top 3 bases that produce the least residual errors when used in the reconstruction of the original image. These features are observed to preserve the structure of the original image and therefore can be mapped onto the RGB of a colour image as shown in Figure 2

## 2.2 Generating Superpixels

Aggregating the pixel features into superpixels eliminates the local neighbourhood redundancy in the feature space and reduces the computational cost of partitioning the image into the required regions. We perform an oversegmentation of the phase contrast image using simple linear iterative clustering (SLIC) superpixels [1]. For an image with $N$ pixels, we obtain $K$ superpixels which consist of small regions with an approximate size $N/K$ that stay within the object boundaries in the original image as shown in Figure 2. Hence, for roughly equally-sized superpixels there would be a superpixel center at every grid interval $Q = \sqrt{N/K}$.

For an image with phase retardation features $\{\Psi_x, \Psi_y\}$ for arbitrary pixel locations $\{x, y\}$, the similarity distance measure for aggregating the pixels, $d_s$ is defined as:

$$d_s(x,y) := (\Psi_x - \Psi_y) + \frac{c}{Q}||x-y||_2 \tag{2}$$

where $d_s$ accounts for both feature variation and distance in pixel location. The compactness, $c$ determines the extent of emphasis on spatial proximity. The higher its value, the more compact the superpixels.

## 2.3 Multiclass Clustering and Region Extraction

We perform multi-class clustering on the features associated with each superpixel to group the segments into regions that correspond to the main region partitions expected in a typical phase contrast image. These include dark cells, bright cells, and halos.

Each superpixel is initially characterized by the mean and covariance of the multi dimensional pixel phase feature vector within it. It is observed that bright cells and halos are both represented by similar phase retardation features (shown in red in Figure 2(b)). Hence bright cells and halos cannot be distinguished by using only the features within each superpixel. As these are usually surrounded by different structures, we augment the feature vector of each superpixel by incorporating neighbourhood information. To include neighbourhood information, we construct the adjacency graph, $G(V,E)$, of the superpixels $s_i \in S$, where $E$ is the set of edges formed between pairs of adjacent superpixels $(s_i, s_j)$ in the image.

By estimating the dissimilarity between the feature distribution, $P_s$, in a given superpixel, $s$, and the feature distribution, $P_{N(s)}$, in its neighbourhood, $N(s)$, we capture some basic context information. We use the Jensen-Shannon divergence (JSD) [4], which is a symmetrized and smoothed version of the Kullback-Leibler divergence, $KL$, can be written as:

$$JSD(P_s||P_{N(s)}) = \frac{1}{2}KL(P_s||M) + \frac{1}{2}KL(M||P_{N(s)}) \quad \text{where} \quad M = \frac{1}{2}(P_s + P_{N(s)}) \,. \tag{3}$$

The divergence between $P_s$ and $P_{N(s)}$ is used to select the most distinctive neighbour feature which is used as an additional feature for each superpixel $s$. As a result, we obtain a feature vector for each superpixel that consists of its mean and covariance, and that of its most distinctive neighbour. This ensures that neighbourhood information that is included is the one that shows the most deviation from the distribution of a given superpixel.

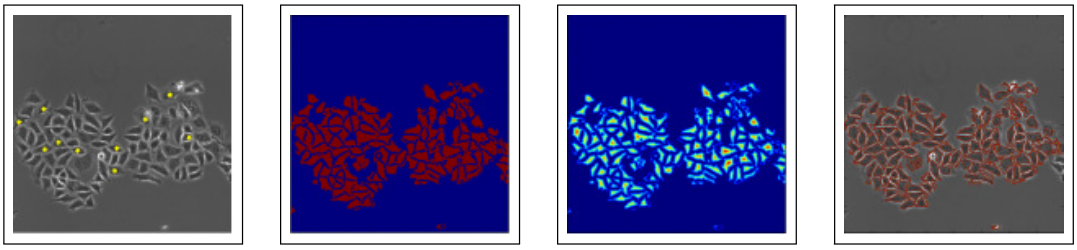
Considering every superpixel is modelled by the mean, $\mu$ and covariance, $\sigma$ of a multivariate Gaussian, we proceed to make the assumption that groups of superpixel features that share similar mean and covariance could be modelled as the components of a Gaussian Mixture Model (GMM). We then proceed to fit the superfeatures to $q$ components of a GMM via expectation maximization (EM). This amounts to clustering the superpixels in a given image into $q$ clusters. Figure 2(d) shows the output of the clustering.

We observe that the clusters partition the image into various regions. Considering this partitioning is completely unsupervised, we proceed to to extract the label associated with cell regions by incorporating prior knowledge of some pixel coordinates indicating cell location. A single user mouse click in any part of a cell region corresponding to a coordinate per cell is required for only a few cells in the image (See Figure 3(a)). We obtain a binary mask, as shown in Figure 3, of cell regions and background for each image by labelling all regions that have the same label as the known cell pixel points to 1 (foreground) and all other regions to 0 (background).

We further observe that some cell regions (especially for cluttered cells) in the mask merge along some length of their shared boundary. To split such regions, we compute the Euclidean distance transform on the mask such that the value of each pixel is replaced by its distance to the nearest background pixel (See Figure 3). We then find the local peaks in the distance transform and label them as markers. We the proceed to run a marker-controlled watershed routine to find the *watershed line* that separates adjoining regions.

# 3 Experiments and Results

10 phase contrast images of cervical cancer cell colonies of the HeLa cells were used to evaluate the proposed algorithm. The cells contained in the data set vary significantly in size and shape. Each image is $400 \times 400$ pixels in size and contains roughly 100 cells each imaged at a magnification of 10X.



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

Figure 3: Final processing leading to final segmentation output: (a) *a priori* knowledge of cell location as dots (b) Cell region mask; (c) Euclidean distance transform map; (d) Segmentation contours overlaid on image

## 3.1 Parameter Selection

We detail the various parameter settings we adopted for the experiments run on our data here. In computing the phase retardation features for each pixel, we set the length of the dictionary of diffraction patterns, $M = 20$ resulting in a 20-dimensional feature vector for each pixel. Also, we set the number of superpixels, $K = 2500$ for each image with a corresponding high compactness, $c = 20$ to emphasize spatial compactness that ensures that the superpixels do not spread over boundaries in the image. For a $400 \times 400$ image with $N = 160000$ pixels, we obtain an average superpixel size, $N/K = 64$ pixels and grid size $Q = 8$. For multiclass clustering using GMM, we set $q = 4$ components representing dark cell regions, bright cell regions, halos and background.

## 3.2 Evaluation

We perform a quantitative assessment of our method on our data by computing the pixel-wise Tanimoto coefficient (TC) introduced in [5] as:

$$TC = \frac{L \cap L_o}{L \cup L_o} \tag{4}$$

where $L$ is the predicted label obtained from the final segmentation mask and $L_o$ is the ground truth label obtained by manually segmenting all cells in all images used.

We achieve an average TC of $0.94 \pm 0.03$. Table 1 compares our result to that of similar work. We achieve similar performance to Su *et al.*[7] yet provide only weak cell location annotation with dots for only about 5% of cells in an image compared to 20% of complete cell segmentation in [7].

Our approach however is dependent on majority of the labelled positions being correctly grouped into the cluster that corresponds to cells. In the event of poor clustering, there is a risk of poor delineation of boundaries when the majority of extracted labels from a priori locations in the multiclass clustering do not correspond to that cells regions. This is however unlikely to occur in our case unless the features are poorly extracted and the clustering fails completely.

| Method | Tanimoto Coefficient |
|---|---|
| Our proposal | $0.94 \pm 0.03$ |
| Su *et al.* [7] | $0.94 \pm 0.02$ |
| Yin *et al.* [9] | $0.83 \pm 0.06$ |

Table 1: Quantitative comparison with similar work

# 4 Conclusions and Further Work

We have demonstrated a method for cell segmentation by utilizing phase retardation features that compares in performance to the state-of-the-art. It should be considered that in comparison to Su *et al.* [7] we make use of weaker and fewer annotations. While annotations are typically a very good way for improving algorithm performance, it is typically hard to obtain these in practice. Further work is needed to limit the number of parameter adjustments that are necessary to account for different imaging conditions, especially in case intrinsic values of the microscope are not known.

We aim to utilize the proposed segmentation algorithm for tracking cells in time-lapse imagery. The only annotation we will require is to obtain the cell location at the time of initialization. But it would also be possible to obtain this information using some type of purpose build object detector. For tracking cells in any consecutive frame, the intrinsic location information from the tracking process could be exploited to increase the amount of prior location knowledge and hence improve the robustness of the overall segmentation. Cell state parameters based on the cellular morphology from such segmentation and tracking in time-lapse could then be obtained efficiently.

Furthermore, we are interested in extracting additional parameters that characterize each cell. Here it is, for example, possible to extract features on the nuclei regions and treat these as features. Additional studies are necessary to analyse if changes in nuclear texture can be correlated with functional fluorescent markers.

# References

[1] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Fua Pascal, and Susstrunk Sabine. Slic superpixels. *EPFL Technical Report 149300*, (June), 2010.

[2] Rehan Ali, Mark Gooding, Tünde Szilágyi, Borivoj Vojnovic, Martin Christlieb, and Michael Brady. Automatic segmentation of adherent biological cell boundaries and nuclei from brightfield microscopy images. *Machine Vision and Applications*, 23(4): 607–621, May 2011.

[3] Carlos Arteta, Victor Lempitsky, J Alison Noble, and Andrew Zisserman. Learning to detect cells using non-overlapping extremal regions. *MICCAI*, 15:348–56, January 2012.

[4] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, 1991. ISBN 0-471-06259-6.

[5] W.R. Crum, O. Camara, and D. L G Hill. Generalized overlap measures for evaluation and validation in medical image analysis. *Medical Imaging, IEEE Transactions on*, 25 (11):1451–1461, Nov 2006.

[6] David J Stephens and Victoria J Allan. Light microscopy techniques for live cell imaging. *Science (New York, N.Y.)*, 300(5616):82–6, April 2003.

[7] Hang Su, Zhaozheng Yin, Seungil Huh, and Takeo Kanade. Cell segmentation in phase contrast microscopy images via semi-supervised classification over optics-related features. *Medical image analysis*, 17(7):746–65, October 2013.

[8] Xiaoxu Wang, Weijun He, Dimitris Metaxas, Robin Mathew, and Eileen White. Cell segmentation for division rate estimation in computerized video time-lapse microscopy. In *Proceedings of the SPIE*, 2007.

[9] Zhaozheng Yin, Takeo Kanade, and Mei Chen. Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation. *Medical image analysis*, 16 (5):1047–62, July 2012.

# Automated Vessel Boundary Detection Using 3D Expansion of Dynamic Programming

Da-Chuan Cheng* Prof
http://mrt.cmu.edu.tw/english/faculty.html
Department of Biomedical Imaging and Radiological Science
China Medical University
Taichung, TW

Shing-Hong Liu Prof
http://csie.cyut.edu.tw/~shliu/
Department of Computer Science and Information Engineering
Chaoyang University of Technology
Taichung, TW

**Abstract**

In this paper we apply the 3D expansion of **DP** (dynamic programming) to find a near optimal surface in a 3D matrix. This algorithm can detect boundaries in an image sequence such as B-mode sonographic images and MRI sequences. We test and modify this algorithm using real B-mode **CCA** (common carotid artery) dynamic sonographic image sequences and **FA** (femoral artery) MRI sequences. The automated results are compared with experts' manual tracing results. The average error of intima detection is around 0.5 pixel (two sequences, 164 total images) and the relative error rate for vessel cross-sectional area computation in MRI is $2.8\% \pm 2.2\%$ on 6 MRI sequences (300 total images). The results demonstrate the better qualitative performance than the 2D DP.

## 1 Introduction

Many image edge detection methods have been proposed in the last two decades [5, 6]. Edge detection methods allow broken or discontinuous edge lines. Boundary detections, however, are somehow different, as some applications do not allow discontinuous boundaries. Many types of methods can handle the vessel extraction problem. Several hundred papers are surveyed in [3], which consider vessel extraction mainly in two-dimensional images. DP is an optimization tool often used in boundary detection [4]. The motivation of this study is raised from the observation that the femoral artery in the cross-sectional view in MRI sequences without contrast injection is sometimes invisible. The artery is visible if the blood flow velocity is large enough that it appears in contrast compared to the surrounding tissue. If the blood flow velocity is small during a short time period, the artery has limited contrast and is not visible. Under this extreme situation, the 2D DP fails to detect its boundary because it is hard to obtain a feature, normally the gray-level gradient, to represent the boundary.

To solve this problem, we have found a literature proposed by [2], which is a 3D-expansion of DP that can seek an optimal surface in a 3D matrix. The significance of this

©2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

paper is two-fold: 1) we modify the cost function in the 3D-expansion of DP so that this algorithm is able to find a closed contour and garantees the start and end points are connected; 2) we test this algorithm on the vessel boundary detection on real medical image sequences, and its unsigned relative error is reported. The results show that the proposed method is qualitatively better than the traditional 2D DP.

# 2 Methods

## 2.1 2D dynamic programming

DP is suitable for boundary detection in noisy images, including sonography [1, 4] and MRI vessel segmentation without a contrast injection. However, the drawback of 2D DP is that it does not consider the relatioship between 2 subsequent images. Therefore, if one or two images in an image sequence were blurred, 2D DP fails to detect correct boundaries in these blurred images.

## 2.2 3D expansion of dynamic programming for closed contour detection

The real 3D DP needs tremendous computations. This is however not practical. The main idea of 3D expansion of DP is to use some constraints to limit the computation time. At the same time, the continuities in 3D are still considered. However, the original paper [2] does not consider the connection between the start and end points if it processes the closed contour. In this paper, we modify this drawback and add an extra term in the cost function to fix this problem.

Let the 3D matrix R have size $M \times N \times P$, where M and N are the numbers of rows and columns, and P denotes its depth. The volumetric matrix contains feature image sequences of interest. Values in R are normalized, i.e., $0 \leq R(y,x,z) \leq 1$ where $y$, $x$, and $z$ are indices of the corresponding dimensions. Assume that features are saved in the 3D matrix and the goal is to find an optimal surface having the shortest path and lowest value summation from one side on the $z$-axis to another side with some given constraints. The typical constraint in DP controls the smoothness or continuity of the sought surface. Assume the searching direction is from $x = 1$ to $x = N$ and $z = 1$ to $z = P$. Two parameters control the smoothness: $d_1 \geq |y_x - y_{x-1}|$ controls the smoothness on the $x-y$ plane $N \times M$, and $d_2 \geq |y_z - y_{z-1}|$ controls the smoothness on the $x-z$ plane $N \times P$. The parameters allow the maximum distance between two connected nodes. The $y$-coordinate $(y(x,z))$ denotes the height of the optimal surface, which can be found via an optimization problem:

$$y^*_{x,z} = \arg\min_{y_{x,z}} \sum_{x=1}^{N} \sum_{z=1}^{P} \{R(y_{x,z}) | 1 \leq x \leq N,\ 1 \leq z \leq P\}, \tag{1}$$

subject to some constraints $d_1 \geq |y_x - y_{x-1}|$ and $d_2 \geq |y_z - y_{z-1}|$;
This optimization problem minimizes the iterative cost functions as such:

$$C_1(y,x,z) = \min_{i \in (-d_1, d_1)} (C_1(y+i, x-1, z) + \beta_1 |i| + \frac{\gamma |y+i-y_0|}{N-x+1}) + R(y,x,z), \tag{2}$$

for $1 \leq x \leq N,\ 1 \leq z \leq P$. The term with a weighting factor $\gamma$ is added to modify the algorithm so that it garantees the connection between the start and end point if this algorithm processes

the closed contour. The weighting factor $\gamma$ can be set to 1 for closed contour detection and $\gamma = 0$ for other cases. After $C_1$ is fully constructed, $C_2$ is constructed afterward as follows:

$$C_2(y,x,z) = \min_{i\in(-d_2,d_2)}(C_2(y+i,x,z-1)+\beta_2|i|)+C_1(y,x,z), \tag{3}$$

for $1 \le z \le P,\ 1 \le x \le N$;
where $C_1$ and $C_2$ are accumulation matrices of size $M \times N \times P$. The initializations of $C_1$ and $C_2$ are thus:

$$\begin{aligned} C_1(*,1,z) &= R(*,1,z) && \text{for } 1 \le z \le P; \\ C_2(*,x,1) &= C_1(*,x,1) && \text{for } 1 \le x \le N. \end{aligned}$$

Notably, the accumulation matrices $C_1$ and $C_2$ are accumulated on $y$-$x$ plane and $y$-$z$ plane, respectively. The matrix $C_1$ must be done before $C_2$ starts to be computed. We found that [2] does not explain how to preserve the continuity in the $y$-$x$ plane. In order to preserve the constraint $d_1 \ge |y_x - y_{x-1}|$, we make a modification: another 3D matrix $C_3$ is built. The matrix $C_3$ is built backwards from the last $y$-$x$ plane ($z = P$) to the first $y$-$x$ plane ($z = 1$) and the near optimal surface is obtained line-by-line sequentially. The iterative cost function $C_3$ is defined as follows:

$$C_3(y,x,z) = \min_{j\in(-d,d)}(C_3(y+i,x-1,z)+\beta_1|i|+\frac{\gamma|y+i-y_0|}{N-x+1})+C_2(y,x,z), \tag{4}$$

subjected to $z = P, P-1, \cdots, 1$. The optimal index $i^*$ can be determined by the following equation:

$$i^* = \arg\min_{i\in(-d_1,d_1)}(C_3(y+i,x-1,z)+\beta_1|i|+\frac{\gamma|y+i-y_0|}{N-x+1}), \tag{5}$$

Similarly, the optimal index is stored in a 3D matrix: $Y(y,x,z) = y + i^*$, which is a pointer indicating the node on its previous column in the $y$-$x$ plane.

After the process in Eq.(4 and 5) is fully done, the line is obtained via the following process. The last line (z=P) is obtained first:

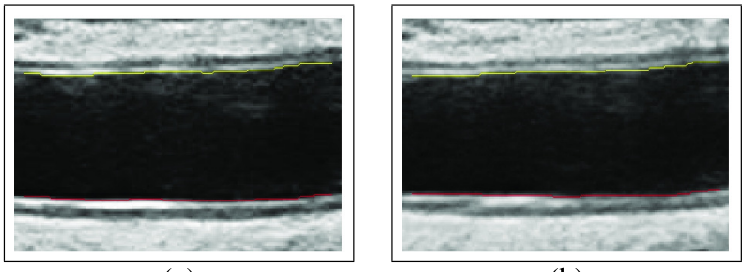
$$\hat{Y}(x,z) = \hat{y_z}, \text{if } z = P;$$

where $\hat{y_z} = \{y_1, y_2, \cdots, y_N\}_{z=P}$ is the line on the $z$-th $y$-$x$ plane obtained by the backwards tracing on $Y(y,x,z)$, subject to $\hat{y}_z = \arg\min_y \sum_{x=1}^{N} C_2(y,x,z)$. In fact, the line $\hat{y}_{z=P}$ is obtained by searching the position where the $C_3(y^*,N,P)$ has the minimal value among the last ($N$-th) column of $C_3$ and then backwards tracing using the 3D matrix $Y(y,x,z)$ from the start point $(y^*,N,P)$.

To retrieve the rest lines ($z < P$), the line is obtained by:

$$\hat{Y}(x,z) = \hat{y_z}, \text{if } 1 \le z < P;$$

where $\hat{y_z} = \{y_1, y_2, \cdots, y_N\}_{1\le z<P}$ is the line on the $z$-th $y$-$x$ plane obtained by the backwards tracing on $Y(y,x,z)$, subject to $\hat{y}_z = \arg\min_{y\in\Omega} \sum_{x=1}^{N} C_2(y,x,z)$, where $\Omega$ defines the condition to satisfy: $\max|\hat{y}_z - \hat{y}_{z+1}| < d_2$. Via this structure, the near optimal surface can be obtained efficiently and the resultant $y$-coordinate are stored in a 2D matrix $\hat{Y}(x,z)$.



(a) (b)

Figure 1: The intima detection of near (upper line) and far walls of CCA in B-mode sonographic image sequences (S1). (Parameter setting: $\beta_1 = 0.1$, $\beta_2 = 0.05, \gamma = 0$, $d_1 = 1$, and $d_2 = 5$).

## 2.3 System error measurement

To estimate the error of the proposed system, two error measurements are defined as follows. They are applied for (1) CCA intima detection on B-mode sonographic image sequences (Eq.(6)) and (2) cross-sectional area of FA in MRI sequences (Eq.(7)).

Assume that the intima coordinates made by the proposed method are stored in a 2D matrix $\hat{Y}(x,z)$, $1 \le x \le N$ and $1 \le z \le P$. The manual drawing performed by experts are denoted by $\tilde{Y}(x,z)$, therefore, the error of the $z$-th line is defined by:

$$\varepsilon_z = \frac{1}{N} \sum_{x=1}^{N} |\hat{Y}(x,z) - \tilde{Y}(x,z)| \tag{6}$$

Assume that the cross-sectional area of the FA measured by the proposed method on the $z$-th image is denoted by $A_z$, and the manual measurement by experts is represented by $\tilde{A}_z$. Then the relative error rate (total N images) is defined by:

$$\xi = \frac{1}{N} \sum_{z=1}^{N} \frac{|Y(z) - \tilde{Y}(z)|}{\tilde{Y}(z)} \times 100 \tag{7}$$

# 3 Results

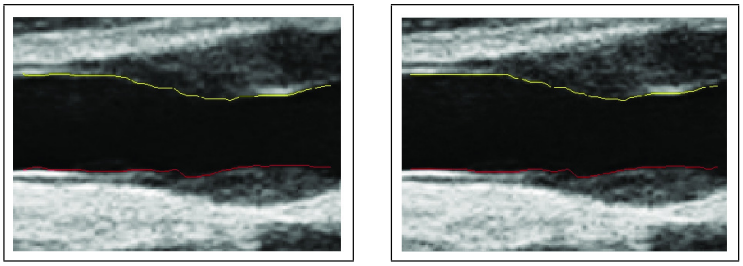
Two B-mode sonographic sequences are used to test. The resultant detected intima lines are shown in Fig. 1. This sequence has a larger IMT (intima-media thickness). Normally the age is a factor to influence the IMT. When IMT is increased the risk getting stroke is also increased. IMT is a clinical index in cardiovascular diseases.

Figure 2 shows a patient having a plaque in CCA. This result demonstrates the ability of the proposed system to detect intima having a plaque. The stenosis resulted from a plaque is a high risk, which causes the blood flow in a unstable way and causes turbulence.

In order to compare the manual tracing and the automated detection results, the Bland-Altman plot is modified.

The average errors (in pixel) are given in Table 1. They are around 0.5 pixel. The proposed method performs very well in detecting the intima in B-mode sonography.

We furthermore apply the method to detect the FA in MRI sequences. Due to page limitation, we have to ignore the description of MRI protocal. Table 2 shows the vessel boundary detection results of six MRI sequences. The average error rate is $2.8\% \pm 2.2\%$. All error rates are under 5%, which is acceptable for clinic measurements.



(a) (b)

Figure 2: The intima detection of near (upper line) and far walls of CCA in B-mode sonographic image sequences (S2). (Parameter setting: $\beta_1 = 0.1$, $\beta_2 = 0.05$, $\gamma = 0$, $d_1 = 1$, and $d_2 = 5$).

Table 1: The average errors in pixel.

| Sequence name | Wall | Average error (pixel):Eq.(6) | Quantity of image |
|---|---|---|---|
| S1 | Near wall | 0.63 | 86 |
| | Far wall | 0.54 | |
| S2 | Near wall | 0.41 | 78 |
| | Far wall | 0.54 | |

Table 2: The average error rate in percentage.

| Sequence No. | Mean(%) | SD(%) |
|---|---|---|
| S1 | 2.4 | 1.8 |
| S2 | 3.2 | 2.1 |
| S3 | 1.4 | 1.3 |
| S4 | 3.2 | 2.7 |
| S5 | 3.6 | 2.6 |
| S6 | 3.1 | 2.4 |

# 4 Discussions

The parameter for $d_1$ is normally set to 1, which implies that the boundary is smooth and no jump in an image. However, $d_2$ can be larger than $d_1$, since it allows the boundary to jump from image to image. This will happen if the artery wall moves fast especially in the dynamic B-mode sonography of CCA. However, the user does not know how much the artery wall moves, therefore, it is crucial to determine $d_2$. If $d_2$ is set too large, it might increase the computation time and more seriously is to affect the result since it might detect the wrong wall by jumping into the neighboring wall.

Our future goal is to improve this algorithm to be able to detect dual surfaces in a 3D matrix, so that the dual lines can be detected simultaneously in processing an image sequence, such as the application on the intima and adventitia detection on CCA in B-mode sonography.

# 5 Conclusion

We provide a modificayion on 3D DP, which garantees the connection of a closed contour. This was not noticed by the original author. Our modification performs qualitatively better than the previously published method on detection of closed contour such as the cross-sectional view of vessels either in MRI or sonography. Based on the experiments to detect the vessel boundaries in real medical images (B-mode sonography and MRI), we conclude that the propose method is able to replace the experts' manual work in vessel boundary detection.

# References

[1] D.C. Cheng and Xiaoyi Jiang. Detections of arterial wall in sonographic artery images using dual dynamic programming. *IEEE Trans. on Information Technology in Biomedicine*, 12(6):792–799, 2008.

[2] D.C. Cheng and Jui-Teng Lin. Three-dimensional expansion of a dynamic programming method for boundary detection and its application to sequential magnetic resonance imaging (mri). *Sensors*, 12:5195–5211, 2012.

[3] C. Kirbas and F. Quek. Vessel extraction techniques and algorithms: a survey. In *Third IEEE Symposium on Bioinformatics and Bioengineering*, 2003.

[4] Q. Liang, I. Wendelhag, J. Wikstrand, and T. Gustavsson. A multiscale dynamic programming procedure for boundary detection in ultrasonic artery images. *IEEE Trans. on Medical Imaging*, 19:127–142, 2000.

[5] L. Zhai, S. Dong, and H. Ma. Recent methods and applications on image edge detection. In *Proceedings of the 2008 international Workshop on Education Technology and Training, 2008 international Workshop on Geoscience and Remote Sensing*, 2008.

[6] D. Ziou and S. Tabbone. Edge detection techniques: An overview. *International Journal of Pattern Recognition and Image Analysis*, 8(4):537–559, 1998.

# Non-Contact Pulmonary Functional Testing through an improved Photometric Stereo Approach

Jahanzeb Ahmad
jahanzeb.ahmad@uwe.ac.uk
Jiuai Sun
jiuai2.sun@uwe.ac.uk
Lyndon Smith
lyndon.smith@uwe.ac.uk
Melvyn Smith
melvyn.smith@uwe.ac.uk

Center of Machine Vision,
Bristol Robotics Laboratory,
University of the West of England,
Bristol, UK

**Abstract**

A non-contact computer vision based system is developed for Pulmonary Functional Testing. The unique and novel features of the system are that it views the patients from both front and back and creates a 3D structure of the whole torso. By observing the 3D structure of the torso over time, the amount of air inhaled and exhaled is estimated. The Photometric Stereo method is used to recover local 3D surface orientation. This technique is good for recovering the local 3D surface orientation but provides no absolute depth information about the observed scene. To calculate absolute depth information from Photometric Stereo a new Lambertian Diffused Maxima Region (DMR) based algorithm is employed here. A high (0.96) correlation between breathing volume calculated from Photometric Stereo and Spirometer data is observed.

## 1 Introduction

The most common test performed on patients to assess their respiratory functionality is known as the Pulmonary Functionality Test (PFT). This test measures the flow and volume of air inspired and expired from lungs using device called a Spirometer. The contact based nature of the devices increases breath awareness which may cause hyperventilation and alter the breathing pattern [5, 6]. In addition, a high degree of compliance is required from patient which is not always possible because of the contact based nature of Spirometer. Recent advancements in computer vision and computing have enabled researchers to develop non-invasive non-contact systems to overcome the problem due to the contact based nature of Spirometers .

The pioneering work in development of a vision based Spirometer was done by Ferrigno and Pedotti [4] in 1985 and their system was called The "ELITE" (ELITE System; BTS, Milan, Italy). Ferrigno and Pedotti placed predefined markers on the patient's torso and used a digital video motion analyser system to track markers. Aoki et al [1] in 2005 presented a

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

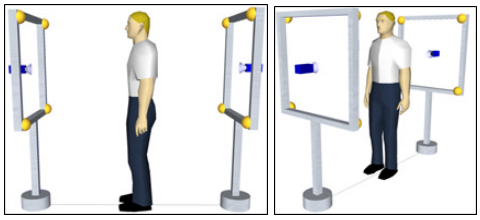


Figure 1: Photometric stereo Rig with subject standing in middle of two rigs.

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

Figure 2: Raw Images obtained from FrontRig. Image(a) when top-right light is turned on, image(b) when bottom-right light is turned on, image(c) when bottom-left light is turned on, image(d) when top-left light is turned on and image(e) when no light is turned on.

triangulation based respiration monitoring and measurement method. The system was based on a fibre grating vision sensor, which is composed of a multiple slit light projector and a CCD camera. Recently a group of researchers from Cambridge University, Cambridge UK and PneumaCare Limited, Duxford UK have developed a new Non-Contact (marker-less) system called Structured Light Plethysmography(SLP) for respiratory functional testing [2].

In this paper we present a novel Photometric Stereo [7] based method for performing pulmonary functional testing through surface gradient analysis. Photometric stereo (PS) is able to recover the surface shape of an object or scene by means of several images taken from a same view point but under different lighting conditions. At least three images are needed to estimate the surface gradient using Photometric Stereo. In our setup we use five images to estimate surface gradient; four images are taken by turning on one light source at a time as shown in Figure 2(a) to Figure 2(d). Light sources are placed at equal distance from each other to get homogenous light distribution among all four images as shown in Figure 1. The fifth image is taken when no light source is turned on to get the background light intensity (noise) as shown in Figure 2(e). As the images are taken in an unconstrained office room environment, there is a fluorescent tube light on ceiling is clearly visible in Figure 2.

## 2 Surface gradient and volume estimation

According to the Lambertian reflectance model the intensity I of light reflected from an object's surface is dependent on the surface albedo $\rho$ and the cosine of the angle of the incident light as described in Equation 1. The cosine of the incident angle can also be referred as dot product of the unit vector of the surface normal N and the unit vector of light source direction $\vec{L}$, as shown in Equation 2. When more than two images from same view point are available under different light conditions, we have a linear set of equations 1 and 2 and this

can be represented in vector form as shown in equation 3.

$$I = cos(\phi_i) \tag{1}$$

$$I = (L \cdot N) \tag{2}$$

$$\vec{I}(x,y) = (x,y)[L]\vec{N}(x,y) \tag{3}$$

In our case $\vec{I}$ is the vector formed by the four pixels $((I^1(x,y), I^2(x,y), I^3(x,y), I^4(x,y))$ from four images, [L] is the matrix composed by the light vectors $(\vec{L^1}, \vec{L^2}, \vec{L^3}, \vec{L^4})$. 1, 2, 3 and 4 is the number with respect to the individual light source direction and (x,y) is pixel position in spatial domain. These light directions are calculated during a calibration process using a specular sphere. [L] is not a square and so not invertible, but the local surface gradients p(x,y) and q(x,y), and the local surface normal N(x,y) can be calculated from the least square method using equations 4 and 5.

$$\vec{M}(x,y) = ([L]^T[L])^{-1}[L]^T\vec{I}(x,y) = (m_1(x,y), m_2(x,y), m_3(x,y)) \tag{4}$$

$$p(x,y) = \frac{-m_1(x,y)}{m_3(x,y)}, q(x,y) = \frac{-m_2(x,y)}{m_3(x,y)} \tag{5}$$

The volume of the torso is represented as the total average gradient; we call it the gradient volume (GV). GV is calculated by adding the absolute average of p and q and then normalizing it as shown in equation 6 and equation 7 for both front and back. Where k and l are the total number of valid pixels in the front and back images respectively, $p_f$,$q_f$ are the local surface gradients of front image and $p_b$,$q_b$ are the local surface gradients of the back image. To compare GV with volume obtained from a Pneumotachometer, normalization is applied on both signals using equation 7. $\bar{V}$ represents mean value of V.

$$V = [\frac{1}{k}\sum(|p_f| + |q_f|)] + [\frac{1}{l}\sum(|p_b| + |q_b|)] \tag{6}$$

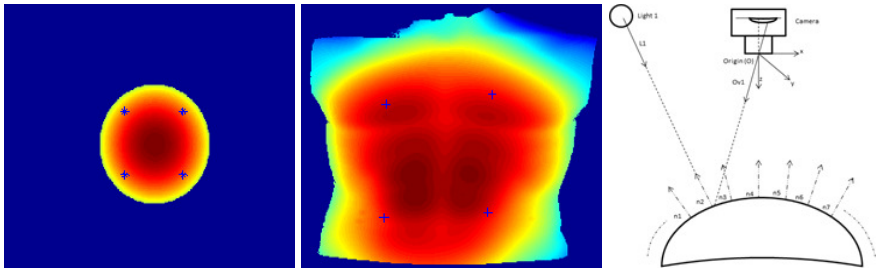
$$GV = \frac{(V - \bar{V}}{max(|V - \bar{V}|)} \tag{7}$$

# 3 Object distance estimation

The object distance from the camera is estimated by using the Diffused Maxima Region (DMR), which is calculated by taking an absolute of dot product between the pseudo light vector and pseudo surface normal and then applying a threshold (th) as shown in equation 8.

$$DMR_i = |N \cdot S^i| > th \tag{8}$$

$S^i$ is a pseudo light vector for light i and N is the pseudo surface normal at each pixel. Pseudo light vector $(S^i)$ is calculated during the calibration process by placing the sphere at the centre of the field of view, it is assumed to be the same for every pixel. The estimated centre of the DMR gives us the point where the surface normal and the light vector are approximately aligned. Many DMR(s) can exist on the surface of an object but the region with



(a) (b) (c)

Figure 3: Diffused Maxima Regions centres of synthetic sphere (a) and real human dummy (b). Depth Calculation using DMR and intersection of vector Ov1 and L1 (c).

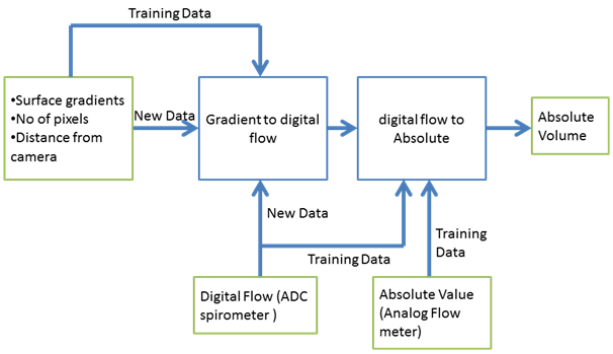



Figure 4: Gradient Volume to Absolute volume conversion.

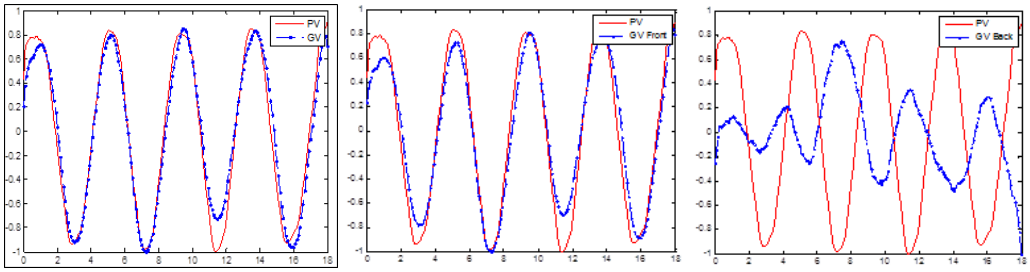
maximum area is considered to be the best choice. DMR can also be selected on the basis of shape or location but experiments have shown that using maximum area as a selection criterion produces best results. Figure 3(a-b) shows the four selected DMR centres on a height map of a synthetic sphere and a real human dummy.

Once the DMR centre is identified in the image plane a vector Ov1 can be created from the DMR centre to the centre of the lens (O) as shown in Figure 3(c). Now by using origin (O), position of light (LP), light vector L1 and vector Ov1 we can determine the intersection point of these two vectors in world coordinates [3]. That effectively gives the knowledge of the scene range from the camera and light sources.

# 4 Calculating absolute amount of air volume

The air volume calculated from photometric stereo in section 2 does not give the absolute amount of air. To calculate absolute amount of air from surface gradient we use two calibration steps. The first calibration step is to convert the gradient information into digital flow value which is achieved through linear regression and this digital flow value is converted to absolute air volume in a second step by using linear curve fitting. The flow of data and calibration process is shown in Figure 4.

The least square method is used for regression, given surface gradient calculated using Photometric Stereo, number of pixels obtained from segmentation and distance of object from camera calculated using diffused maxima, a feature matrix [X] ($X = [gradient\, pixels\, distance]$) is constructed. The signal obtained from a Pneumotachometer is digitized using analog to




(a) (b) (c)

Figure 5: Figure a:Comparison of Pneumotachometer Volume (PV) and Gradient Volume (GV) (combined front and back). Figure b: front Gradient Volume (GV) and Pneumotachometer Volume (PV) plot, Figure c: back GV and PV plot. Horizontal axis is time in seconds. Vertical scale is a normalized representation of volume.

digital converter (ADC) and is represented as pv. $\phi$ ($\phi = (X' * X)^{-1} * X' * pv$) is trained using [X] and pv ($pv_2 = X_2 * \phi$). Where $[X_2]$ is new input and $pv_2$ is the synthesised digital Pneumotachometer signal which will be converted to absolute air volume using curve fitting.

The absolute amount of air inhaled and exhaled is calculated from digital flow by using linear curve fitting. The coefficients of the curve fitting equation are calculated using values obtained from a carefully controlled air flow meter, by digitizing the flow using an analogue to digital converter and transducer. Absolute amount of air is calculated by $absolutevolume = C_1 * pv_2 + C_0$. Where $C_1 = 0.002085, C_0 = -4.007$

# 5 Results

High correlation (0.98) between gradient volume and Pneumotachometer volume (Ground Truth) was found as shown in Figure 5(a). It also reflects the effect of error in surface gradient is very minimal on volume calculation because of the normalization and calibration processes used to calculate the relative and absolute amount of air.

Analysis on the front and back has been performed independently, which suggests that the correlation of the back gradient volume is negatively correlated to the Pneumotachometer volume and front gradient volume is positively correlated to Pneumotachometer volume. Front GV shows high correlation with PV but when back GV, which is weakly and negatively correlated with PV, is added to front GV it yields higher correlation as compared to front GV only. This phenomenon is shown in Figure 5(b-c).

After the calibration process of converting the relative GV to absolute air volume the correlation drops from 0.98 to 0.96. This is due to errors introduced during the training and curve fitting process involved in calibration steps. The experiments have been performed on 6 healthy subjects and results have been summarised in Table 1. Table 1 shows high degree of correlation between air volume obtained from Pneumotachometer and air volume estimated from photometric stereo images.

# 6 Conclusion and future work

We have proposed a novel non-contact approach and designed a device to perform Pulmonary Functional Testing (PFT) on patients of any age. The unique features of the device

Table 1: Correlation between PV and Absolute GV.

| | Correlation between PV and Absolute GV |
|---|---|
| Subject 1 | 0.960 |
| Subject 2 | 0.935 |
| Subject 3 | 0.963 |
| Subject 4 | 0.958 |
| Subject 5 | 0.916 |
| Subject 6 | 0.922 |

are that it needs very little calibration and no registration between front and back view, and generates at least 2000 times higher resolution data compared to similar systems. We also presented a method to calculate the distance of object from a camera using photometric stereo images which gives the absolute amount of air inhaled and exhaled. As the system is based on off the shelf low-cost components it is relatively easy and cheap to make. In future we are planning to use the device for detecting a range of chest and back bone problems and take advantage of the high resolution 3D recovery for regional analysis of the torso.

# References

[1] H. Aoki, K. Koshiji, H. Nakamura, Y. Takemura, and M. Nakajima. Study on respiration monitoring method using near-infrared multiple slit-lights projection. *IEEE International Symposium on Micro-NanoMechatronics and Human Science*, pages 291–296, 2005. doi: 10.1109/MHS.2005.1590006.

[2] Willem de Boer, Joan Lasenby, Jonathan Cameron, Rich Wareham, Shiraz Ahmad, Charlotte Roach, Ward Hills, and Richard Iles. SLP: A Zero-Contact Non-Invasive Method for Pulmonary Function Testing. *British Machine Vision Conference 2010*, pages 85.1–85.12, 2010. doi: 10.5244/C.24.85.

[3] F Dunn and I Parberry. *3D math primer for graphics and game development*. Wordware Publishing, Inc, 2011. ISBN 1556229119.

[4] Giancarlo Ferrigno and Antonio Pedotti. ELITE: a digital dedicated hardware system for movement analysis via real-time TV signal processing. *Biomedical Engineering, IEEE ...*, BME-32(11):943–950, 1985.

[5] R. Gilbert and JH Auchincloss. Changes in tidal volume, frequency, and ventilation induced by their measurement. *Journal of Applied ...*, 33(2):252–254, 1972.

[6] E Schneider, C Dualé, J-L Vaille, L Ouchchane, T Gillart, D Guélon, and P Schoeffler. Comparison of tolerance of facemask vs. mouthpiece for non-invasive ventilation. *Anaesthesia*, 61(1):20–3, January 2006. ISSN 0003-2409. doi: 10.1111/j.1365-2044.2005.04400.x.

[7] RJ Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):139–144, 1980.

# Non-Intrusive Action Symmetry Measurement for Dementia and Rehabilitation

Matthaios Doulgerakis[1]
M.Doulgerakis@kingston.ac.uk
Vasileios Argyriou[1]
Vasileios.Argyriou@kingston.ac.uk
George Tzimiropoulos[2]
gtzimiropoulos@lincoln.ac.uk
Barbara Pierscionek[1]
B.Pierscionek@kingston.ac.uk

[1] Kingston University
London, UK
[2] Lincoln University
London, UK

**Abstract**

The purpose of this work is to quantify the symmetry in human actions, which is a common objective in medical research. To this end, a novel non-intrusive method for quantifying action symmetry in the frequency domain is proposed. Reflectional symmetry is estimated over time and a new metric is proposed based on dynamic symmetry warping distance. The proposed method can be utilised to provide quantitative estimates of symmetry in actions occurring during the examination of a patient. In order to evaluate the proposed approach and metric, experiments were performed both with phantom data where ground truth is available and with real data offering quantitative and qualitative analysis.

## 1 Introduction

The purpose of this work is to quantify the symmetry in human actions, which is a common objective in medical research. The symmetry between the sides of the human body can be a valuable metric for many medical related procedures and applications. For example, monitoring patients with cerebral palsy [9], in rehabilitation after a stroke [2] or amputation [10], diagnosis of Parkinson's disease in early stages [8] and even diagnosis of numerous medical conditions in infants [4].

There are four main categories of symmetry as described in the literature [5], (see figure 1). In research on human body motion analysis, the mirror-symmetry properties are examined, assuming that the axis of symmetry is vertical and the differences between the left and right side motions are considered. In existing literature on human body motion analysis, symmetry has been estimated using a variety of sensors all of which are intrusive to the subject. In particular, to acquire the measurements, a variety of sensors attached to the subjects have been employed (gyroscopes [3], accelerometers [6], motion capture sensors [14]). Additionally, the placement of the sensors and the acquisition of the respective measurements are not universal and they change according to the action or application.

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

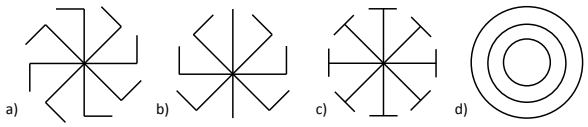



Figure 1: Examples of symmetry (a) rotational, (b) mirror, (c) radial and (d) circular.

To overcome the aforementioned limitations a computer vision approach is proposed to measure symmetry in humans actions. Finding symmetry in images is a well-studied subject. The symmetry detection task is performed either in the pixel [7] or in the frequency domain [12], with the Fourier techniques, that are based on Phase Correlation and log-polar transform to offer measurements with high degree of accuracy and robustness. This approach has the advantage of invariability to changes in illumination, is very fast, scale invariant and can estimate multiple symmetries simultaneously without additional computational cost, [12]. Furthermore, symmetry is an important feature in computer vision and is used for action classification [11].

In this work the concept of symmetry estimation over time with applications in the measurement of the progress of a patient during the rehabilitation or the diagnosis of Parkinson's and Alzheimer's disease is introduced.A novel, non-invasive, method is proposed to quantify symmetry in human actions using RGB or depth sequences. The technique of symmetry detection is extended to pairs of images. The contributions in this research are a novel method to estimate symmetry between a pair of images in the frequency domain and extending this approach to the time dimension by introducing a innovative Dynamic Symmetry Warping technique. The method introduced in this paper can quantify human action symmetry in a variety of cases where symmetry is an important feature for the monitoring of patients and the diagnosis of diseases.

# 2 Methodology

Initially, the subject is asked to perform some activities (e.g. flexion and extension of a limb) both from the left and the right side (see figure 2). The same activity is performed multiple times and then the captured data is pre-processed to extract the action silhouettes. The aim of the proposed algorithm is to estimate the level of symmetry between two actions. Initially symmetry in a single image is defined using properties in the Fourier domain. Next,the approach is extended to estimate the symmetry between a pair of images and finally temporal information is incorporated. As this research is aiming to identify symmetry between the left and the right side of the body, the main symmetry of the body is assumed to lie in the vertical axis of the image plane. The overall process is shown in figure 3.

## 2.1 Estimation of Reflectional Symmetry in the Frequency Domain

Working in the frequency domain allows the robust estimation of translation between images through phase correlation. Furthermore, it has been shown that after converting an image to polar coordinates the rotation can be measured as translation [13]. Exploiting this fact, the method proposed in [12] estimates the symmetry properties of images in the frequency domain, suggesting that the magnitude of the Fourier transform of the polar representation of an image can be used to obtain the symmetry properties of the image. The best approach to achieve a high quality polar representation is to use the fast and accurate polar Fourier transform proposed in [1]. This algorithm's computational complexity is the same order as the standard Cartesian Fast Fourier Transform (FFT), whilst the experimental results show

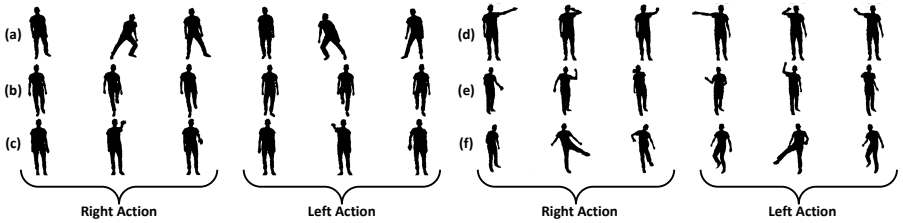



Figure 2: Examples of six symmetry actions performed by the patients in both sides.

high accuracy [1].

The aim is to identify reflectional symmetry in images. To this end symmetry detection is formulated in the frequency domain which enables the use of robust correlation-based techniques. The key advantage of a Fourier-based formulation is that for any type of symmetric pattern about a symmetry centre $(x_0, y_0)$, the magnitude of its Fourier Transform is also symmetric about the origin $(0,0)$. This enables the computation of the axis of symmetry by a simple correlation in the polar domain. In particular, let denote by $I$ a reflectionally symmetric image and let us also denote by $a_0$ the symmetry axis. Based on the Fourier properties, the magnitude $M$ of Fourier Transform of $I$ resampled on polar coordinates $(r, \theta)$ contains the information about the degree and the axis of symmetry that can be obtained by performing a 2D autocorrelation C, as shown below:

$$C = F^{-1}\{M \cdot M^*\} \tag{1}$$

where $M$ is the Fourier transform of the polar magnitude, $*$ denotes the complex conjugate operator and $F^{-1}$ denotes the inverse Fourier transform.

The higher the values in the output signal the higher the symmetry in the corresponding rotation, extrapolating this information allows the detection of the main axes of symmetry of the image as suggested in [12] and as shown can be used to quantify symmetry.

## 2.2 Symmetry Distance Metric

In order to evaluate and quantify the level of symmetry between two given images $I_1$ and $I_2$, a novel metric is proposed. According to the proposed approach, the two images are combined side by side into a larger image $I$, (see figure 3). Concatenating the images in this way ensures that the images found to be symmetrical also have a similar shape. Two dissimilar shapes may have the same symmetry properties but with the undertaken approach the dissimilarity in shape will result in asymmetry. Additionally, the main symmetry is forced to be near to the vertical axis so differences between the sides can be examined, which is the desired result.

After combining the images a 2D autocorrelation is performed on the polar Fourier transformed magnitude of the combined image $I$ (equation (1)). Considering the highest peak in the output signal as the centre, dissimilarities either side suggest asymmetries between the images. The difference between the left $(C_L)$ and the right side $(C_R)$ of the signal is obtained by calculating their distance, as shown in equation (2):

$$S_D = \Delta(C_L, C_R) \tag{2}$$

where $\Delta$ is the Earth mover's distance (EMD) and is calculated by the following equation:

$$\Delta(C_L, C_R) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} d_{ij} f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n} f_{ij}} \tag{3}$$

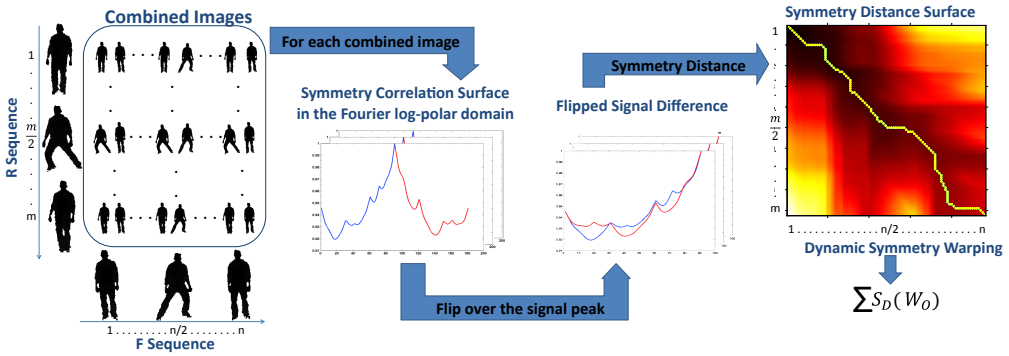



Figure 3: Each of the $m \times n$ combined images is represented in polar coordinates in the frequency domain and the magnitude is used for the 2D autocorrelation. The symmetry distance is calculated as the EMD of the sides of the correlation surface. The symmetry distance surface is composed of the symmetry distance of each combination of images while the index preserves the temporal structure of the original sequences. The yellow line represents the optimal path $(w_o)$ that minimises the global symmetry distance between the two sequences.

where $d$ denotes the ground distance between samples $i$ and $j$ along $C_L$ and $C_R$ and $f$ the flow between them. According to the proposed metric, the smaller the symmetry distance $S_D$ the higher is the symmetry between the images.

## 2.3 Dynamic Symmetry Warping

A symmetry detection and evaluation approach over time is introduced in this section allowing the quantitative assessment of action symmetry. In order to find the symmetry between two time series a novel approach based on the proposed symmetry distance metric is introduced. Dynamic Symmetry Warping (DSW) is a dynamic algorithm that computes the optimal symmetry between two sequences, by stretching the time dimension and summing the local symmetries of individual matched elements. The outputs of the DSW algorithm are the optimal warping of the sequences and the minimum global symmetry distance between the given series.

Specifically, each image of the sequence $R$, of length $m$, is combined with each image of the sequence $F$, of length $n$, and their symmetry distance is calculated as described in equation (2). The $S_D$ values are arranged in an $m \times n$ symmetry distance surface matrix where the indices of $m$ and $n$ describe the temporal structure of $R$ and $F$ respectively. The warping path $w$ that minimises the global symmetry distance $G_W$ between the sequences, indicates the optimal global symmetry distance $G_S(R,F) = \min_w G_w(R,F)$ where $G_w$ are the possible available paths of the accumulated symmetry distance surface:

$$G_w = \sum_{t=1}^{T} S_D(w_R(t)w_F(t)/N_w) \tag{4}$$

where $N_w$ is a normalisation factor and

$$w_R(t) \in \{1...m\} \, and \, w_F(t) \in \{1...n\} \tag{5}$$

where $t = 1...T$, and $T$ is the length of each path.

The normalisation factor $N_w$ can be chosen according to the needs of the application, and determines the flexibility of the path selection. The output $G_S$ is the optimal global symmetry difference between the image sequences after they are stretched in order to have

the best possible symmetry. This metric can be used to assess symmetry in a variety of human actions and to assist the monitoring of patients and the diagnosis of diseases.

# 3 Results

In this section, the experiments performed and the results obtained are presented. In order to evaluate the proposed method, a new Action Symmetry Database (ASDB) was created that consists of both phantom and real data. The phantom dataset is a moving dummy where the main attributes of the motion for each case are known, providing ground truth for each case. The differentiated attributes are the speed and the angle of movement; and the shape of the limbs. Each one of the attributes is subjected to three variations: original state, 10% increment, and 20% increment, resulting in 27 different cases. The second part was captured with the Kinect sensor consisting of 5 subjects performing different actions, including: body weight shifting to each side, arm swings, leg swings, throwing a ball, kicking (see figure 2). Each person performs 6 actions and repeats each action 4 times on the right and 4 times on the left side.

The optimal global symmetry distance $G_S$ metric is used to quantitatively evaluate our algorithm. For the experiments, the actions are considered pre-segmented, with each subject starting in a specific pose, performing an action and finishing in a specific pose. The normalisation factor that was used is $N = m+n$, providing a symmetric continuity constrain forcing each image to be matched at least once, even if that means that some images are matched more than once. In figure 4, the results for the phantom data are displayed, indicating that the proposed metrics are close to ground truth and follow the same trend since the values of our metric are increasing while we differentiate the motion attributes to induce asymmetry. The plots show the response of the system when we keep two parameters constant and increase the third. Also the variance is plotted to provide a more accurate representation of the performance of the proposed metrics. All the results are summarised in table 1.

Table 1: Phantom Data Estimated Mean Error and the Variance.

| Angle Changes | | | Length Changes | | | Speed Changes | | |
|---|---|---|---|---|---|---|---|---|
| 0% | 10% | 20% | 0% | 10% | 20% | 0% | 10% | 20% |
| Length 0% | | | Angle 0% | | | Angle 0% | | |
| 0.336 (0.014) | 1.462 (0.001) | 3.03 (0.008) | 0.336 (0.014) | 2.227 (0.005) | 4.686 (0.001) | 2.345 (1.839) | 2.433 (1.763) | 2.471 (1.740) |
| Length 10% | | | Angle 10% | | | Angle 10% | | |
| 2.227 (0.005) | 3.491 (0.002) | 5.244 (0.012) | 1.462 (0.001) | 3.491 (0.002) | 5.931 (0.005) | 3.652 (1.851) | 3.619 (1.816) | 3.612 (1.814) |
| Length 20% | | | Angle 20% | | | Angle 20% | | |
| 4.686 (0.001) | 5.930 (0.005) | 7.738 (0.014) | 3.03 (0.008) | 5.244 (0.012) | 7.738 (0.014) | 5.448 (1.949) | 5.331 (1.912) | 5.234 (1.908) |

Table 2: Real Data Estimated Mean Error and the Variance.

| Actor | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Symmetry | 6.237 (1.167) | 5.876 (1.563) | 7.308 (1.101) | 8.186 (2.549) | 5.452 (1.562) | 9.399 (3.004) |

Regarding the real data the obtained results are shown in table 2. The more complex action of kicking (action 6) is scoring the highest value indicating that is performed the least symmetrical, as expected. The results show that the overall methodology provides an accurate non-intrusive action symmetry evaluation framework.

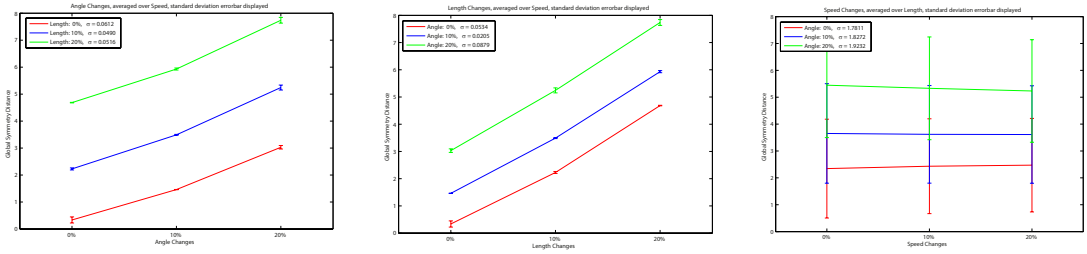


Figure 4: Obtained Action Symmetry Values for the phantom data using the proposed Symmetry Metric.

# 4 Discussion

In this paper a novel framework is proposed to detect and quantify symmetry between actions in a non-intrusive way. The suggested method operates in the frequency domain based on the properties of the Fourier transform. The proposed method has been quantitatively and qualitatively evaluated on the new Action Symmetry Database (ASDB) using both phantom and real samples.

# References


[1] A. Averbuch, P. R. Coifman, D. L. Donoho, M. Elad, and M. Israeli. Fast and accurate Polar Fourier transform. *Appl. Comput. Harmon. Anal.*, 21(2):145–167, 2006.

[2] M. E. Brandstater, H. de Bruin, C. Gowland, and B. M. Clark. Hemiplegic gait: analysis of temporal variables. *Archives of physical medicine and rehabilitation*, 64:583–587, 1983.

[3] S. M. N. A. Senanayake D. Gouwanda. Identifying gait asymmetry using gyroscopes–a cross-correlation and Normalized Symmetry Index approach. *J. Biomech.*, 44(5):972–978, 2011.

[4] P. Venuti G. Esposito. Symmetry in Infancy: Analysis of Motor Development in Autism Spectrum Disorders. *Symmetry (Basel).*, 1(2):215–225, 2009.

[5] Y. Gong, Q. Wang, C. Yang, Y. Gao, and C. Li. Symmetry Detection for Multi-object Using Local Polar Coordinate. *Computer Analysis of Images and Patterns*, 5702:277–284, 2009.

[6] X. Huang, J. M. Mahoney, M. M. Lewis, G. Du, S. J. Piazza, and J. P. Cusumano. Both coordination and symmetry of arm swing are reduced in Parkinson's disease. *Gait Posture*, 35(3):373–377, 2012.

[7] S. Lee and Y. Liu. Curved Glide-Reflection Symmetry Detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(2):266–278, 2012.

[8] M. D. Lewek, R. Poole, J. Johnson, O. Halawa, and X. Huang. Arm swing magnitude and asymmetry during gait in the early stages of Parkinson's disease. *Gait and Posture*, 31:256–260, 2010.

[9] W. L. Minear. Special article: A classification of cerebral palsy. *Pediatrics*, 18(5):841–852, 1956.

[10] R. O. Andres and S. K. Stimmel. Prosthetic alignment effects on gait symmetry: A case study. *Clin. Biomech.*, 5:88–96, 1990.

[11] K. Schindler and L. van Gool. Action snippets: How many frames does human action recognition require? *IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 1–8, 2008.

[12] G. Tzimiropoulos, V. Argyriou, and T. Stathaki. Symmetry detection using frequency domain motion estimation techniques. *Acoust. Speech Signal Process ICASSP*, pages 861–864, 2008.

[13] G. Tzimiropoulos, V. Argyriou, and T. Stathaki. Subpixel Registration With Gradient Correlation. *IEEE Transactions on Image Processing*, 20(6):1761–7, 2011.

[14] R. A. Zifchock, I. Davis, J. Higginson, and T. Royer. The symmetry angle: a novel, robust method of quantifying asymmetry. *Gait Posture*, 27(4):622–627, 2008.

# Analysis of images from optical coherence tomography for the nanoparticle transport

Radim Kolar[1,2]
kolarr@feec.vutbr.cz

Petra Podlipna[1,2]
podlipna@phd.feec.vutbr.cz

1. Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Brno, Czech Republic

2. International Clinical Research Center, Center of Biomedical Engineering, St. Anne's University Hospital, Brno, Czech Republic

**Abstract**

This paper describes an analysis of contrast image sequences from optical coherence tomography. The application of bare nanoparticles in unique single fibre flow phantom is performed in order to examine their spatial and temporal behaviour inside the fibre and the wall of the fibre. The speckle variance method is used to show its usefulness in analysis of contrast OCT sequences.

## 1 Introduction

Optical coherence tomography (OCT) is relatively new imaging modality, which combines microseconds temporal resolution and micron spatial resolution. It enables to study optical properties (local reflectivity) of tissues and other samples with high dynamic range [3]. OCT was introduced in 1991 and since then a number of enhancements, modifications and applications have been presented [2]. The applications cover different fields of clinical, preclinical and experimental research ranging from ophthalmology, intracoronary imaging small animal experiments and nanoparticle research [1, 7]. OCT technique enables to evaluate both, the morphological and functional properties of examined samples/tissue. The functional OCT (fOCT) imaging is an important tool for investigating the processes in tissues of in vivo and in vitro systems. It includes different aspects of tissue/organ function, which can be divided into the contrast imaging approaches using bare nanoparticles, molecular imaging using target nanoparticles and hemodynamic or flow imaging. These categories are of course closely connected.
In this paper, we examine the behaviour of nanoparticles in capillary-simulating phantom via image processing methods. We describe application of nanoparticles for flow visualization and for their transport through semi-permeable membrane in custom-made single-fibre flow phantom.

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

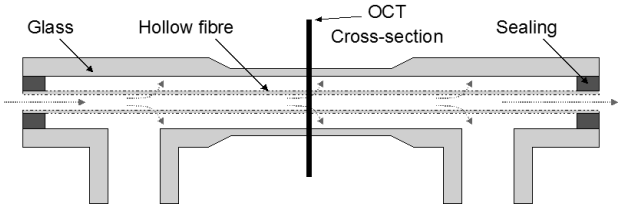



Figure 1: Scheme of the single-fibre flow phantom.

# 2 Method

## 2.1 Data acquisition on single-fibre phantom

In 2013 we presented the construction of a perfusion flow phantom for OCT [6]. Advanced construction of this single fibre phantom has been used for the purpose of this experiment (Figure 1). The polypropylene hollow fibres with pore sizes over $200\,\mu m$ has been used (www.zena-membranes.cz). The fibre was placed in glass capillary with four openings. Two openings were used for setting the inner flow and other two openings for filling the glass capillary with water to ensure optical path without strong reflections. Two syringe pumps (NE-1010, New Era, USA) were used - the first pump (Pump 1) for setting the water flow rate (FR1) through the fibre and the second pump (Pump 2) for constant infusion of the nanoparticles by defined flow rate (FR2). The flow rate of Pump 1 has been constant ($100\,\mu l/min$ ) during the whole experiment. The flow rate of Pump 2 was changing from 100 to $500\,\mu l/min$ with $100\,\mu l/min$ step. Therefore, with increasing FR2 the concentration of nanoparticles increases with ratio given by $FR2/FR1$. For each setting, one OCT sequence has been recorded with framerate 35 fps. The acquisition time was around 10s to ensure that steady state of nanoparticles flow occurs.
In this experiment, gold nanorods (Nanopartz Inc., A12N-25-1400) were used. They have plasmon-resonant peak matching the wavelength of our OCT system (Thorlabs, Swept Source OCT system, OCS1300SS, center wavelength 1325 nm). The diameter of these nanorods is $25\,nm$ and length $256\,nm$. The concentration is $0.05\,mg/ml$.

## 2.2 Data analysis

The data were analysed in three ways:

1. We examined whether the spatial distribution of nanoparticles is homogeneous through the fibre cross-section or not. (Task 1)

2. We analysed whether there is some relation between the amount (concentration) of nanoparticles and image intensity or not. (Task 2)

3. We examined whether the nanoparticles penetrate through the fiber's wall or not. (Task 3)

For this purpose, we used analysis with speckle variance approach, which has been described in [4, 5]. Alternatively, we also examined the temporal changes of image intensity. We also compare these two approaches (speckle based and intensity based).

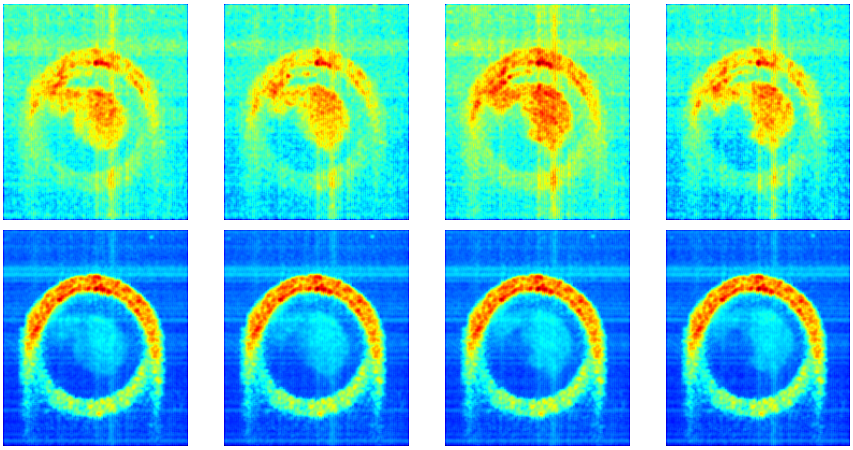


Figure 2: Four SV images (top row) and corresponding intensity images (bottom row) computed from the steady state for $FR2 = 100\,\mu l/min$ in different time intervals. Increased intensity inside the fibre shows presence of nanoparticles.

### 2.3 Speckle variance method

Here we briefly describe principle of the speckle variance (SV) optical coherence tomography method and its adaptation for our analysis. Let assume a sequence of OCT B-mode image $I_i(j,k)$, where $j$ and $k$ are spatial indexes in image matrix and $i$ is temporal index. Inter-frame speckle variance between $N$ images is computed as:

$$SV(j,k) = \frac{1}{N}\sum_{i=1}^{N}\left(I_i(j,k) - I_{mean}(j,k)\right)^2, \tag{1}$$

where $I_{mean}$ is an averaged image obtained from $N$ frames. If we examine the flow or movement of some particles, the speckles differ from frame to frame and SV image thus represents this variations. It is assumed that there is no other movement (e.g. tissue or OCT probe) in examined scene or it is supposed that these tissue displacements between frames are less than the OCT beam radius. It is also supposed that other sources of noise are negligible (e.g. shot noise, electronic noise). The gate length $N$ plays an important role. Authors of [5] used higher values (N=4 or 8) for static scene and the minimum acceptable value (N=2) for scene with possible movements. This value of course depends on current framerate. It is also suggested to averaged $M$ consecutive SV images to improve image quality [5]. $M$ value depends on the properties of the scene. For static scene $M$ value can be large and for dynamic scene smaller or even $M = 1$.
We changed the SV definition above by replacing the quadratic term by absolute value, in order to be able to represent the SV image in intensity levels. We also applied a spatial averaging filter ($3 \times 3$ kernel) for smoothing the final SV image when examining the uniformity of nanoparticles distribution. This helps suppress noisy pixels and also slightly increases the contrast.

## 3 Results and Discussion

*Task 1*
Several SV images computed at steady flow rate in different intervals (1 second) are shown

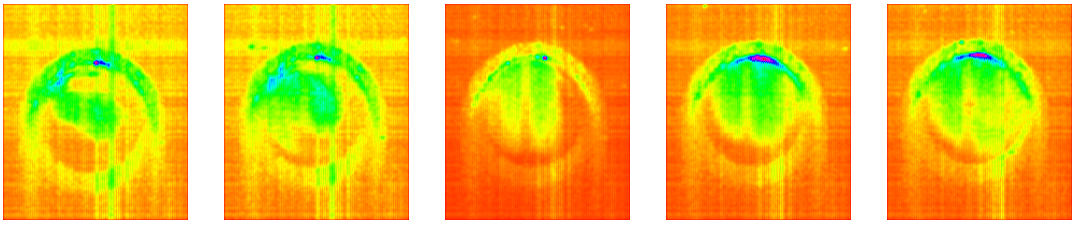


Figure 3: Five SV images computed from the steady state for (from left to right): $FR2 = 100\,\mu l/min$, $FR2 = 200\,\mu l/min$, $FR2 = 300\,\mu l/min$, $FR2 = 400\,\mu l/min$ and $FR2 = 500\,\mu l/min$. Increased intensity inside the fibre shows occurrence of nanoparticles during 3 seconds long interval.

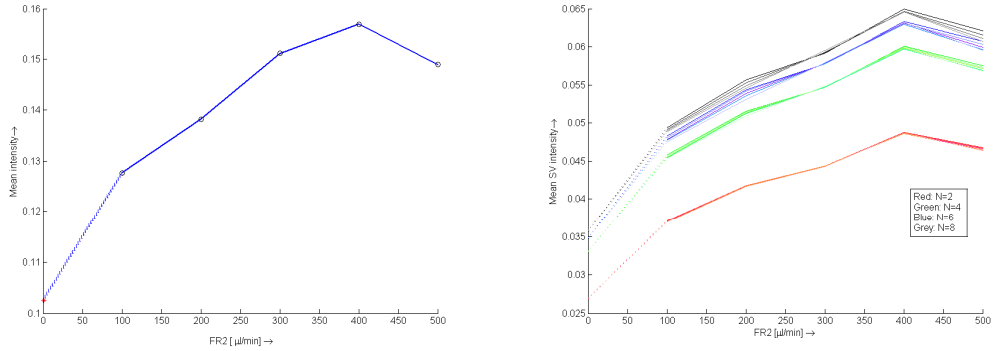
in Figure 2 (top row) for $FR2 = 100\,\mu l/min$. Each image has been created from 16 frames (four frames were used for creating particular SV image and four SV images were averaged). Considering the temporal sampling ($28\,ms$), each image can be examined as a cross-sectional flow during $448\,ms$ interval. The same temporal window was used for creating the averaged intensity sequence (every 16 frames were averaged to create new image), see Figure 2 (bottom row). Highly non-uniform spatial distribution of nanoparticles can be seen in both sequences. It can be also seen that the contrast in the SV sequence was higher than in the average intensity sequence. Similar images can be obtained also for other FR2 values. This is shown in Fig. 3. Each figure represents mean speckle variance computed with $N = 5$ and $M = 21$, which corresponds to 3 seconds interval. It can be observed that as FR2 increases (amount of nanoparticles is increasing and also the total flow rate increases), the nanoparticles flow became more spatially homogeneous.

*Task 2*

In the next step, we analysed relation between the amount of nanoparticles and image intensity for each FR2 values. Firstly we computed mean intensity inside the fibre using manually defined circular mask. The mean value has been computed in both, spatial and temporal domain to obtain one value representing the mean intensity (for intensity sequence) or mean SV value (for SV sequence). To analyse the influence of $N$ and $M$ parameters (see Section 2.2), the SV sequence has been computed for different values of these parameters. Figure 4 shows our results. We can see similar increasing trends for both, intensity and SV sequence. Saturation for high FR2 is clearly seen, which probably means that the concentration of nanoparticles is high and attenuation dominates over scattering. Figure 4b also shows influence of $N$ and $M$ on the result mean SV intensity. We can see that increasing $N$ value increases also SV mean intensity, but the difference became low for $N = 6$ and 8. The multiple colour lines represents different numbers of averaged SV images used for creating the final SV image. We can conclude that this number doesn't significantly influence the SV values. We also computed correlation coefficient between FR2 and intensity or SV values, respectively for FR2 range $100 - 400\,\mu l/min$. All values were above 0.95, which signifies high linear relation between these variables.

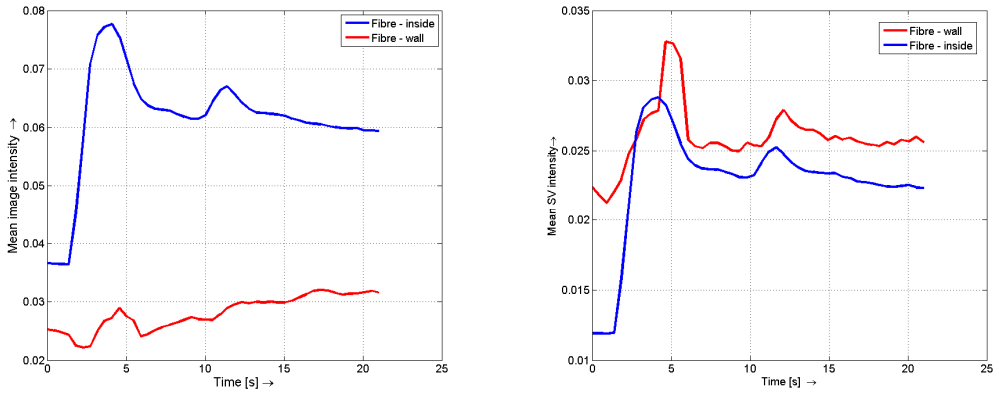
*Task 3*

In the last step of our analysis we evaluated the transport of nanoparticles through semipermeable membrane of the fibre. We evaluated the time course of the intensity and SV value in the whole sequence inside the fibre and within the wall of the fibre in a case of fast application of bolus of nanoparticles. The additional binary annular masks for this analysis have been defined manually.




a) b)

Figure 4: Curves showing the spatial and temporal mean intensity (a) and mean SV intensity (b) computed inside the fibre as a function of FR2. Value for $FR2 = 0$ represents measurement only with water. The x-axis also represents concentration of nanoparticles as described in Section 2.1.




a) b)

Figure 5: Time curves from inside (blue) and from the wall (red) of the fibre in fast bolus experiment for a) image intensity sequence and b) SV sequence.

Figure 5 shows these curves for mean intensity and mean SV sequences. Blue curves (with higher dynamic range) represents change in image intensity or SV intensity, respectively, inside the fibre during and after bolus administration. The first peak represents the maximum concentration achieved during administration of the bolus and the second peak probably represents the end of bolus as there is some pressure distortion in the circuit. It can be seen that the intensity and SV curves extracted from the wall of the fibre are different. The SV curve follows the shape of the curve from inside of the fibre with small delay (around 0.5 second) as can be expected. In the intensity curve, only the first delayed peak can be identified. The dynamic range of this curve is also smaller, which show limited sensitivity. The intensity changes in the fibre is very low, because the light reflections from this environment is high and superimpose the small variations of intensity. It must be noted that the values of this curve has been decreased (because the image intensity inside the wall is significantly higher) to be able to compare both curves.

# 4 Conclusion

We have shown the usability of speckle variance method for analysis of OCT sequences of scene containing nanoparticles. The SV approach achieved higher contrast in resulting images in comparison to intensity analysis (Task 1). SV is also useful for evaluation of the nanoparticle's movement through reflective environment (Task 3). We also observed a linear relation between concentration of nanoparticles in defined range, represented by FR2 (Task 2). It must be noted that SV method requires static scene, which can be difficult to achieve in some applications. However, frame-to-frame registration can be used for sequence stabilization. Another phenomena is signal attenuation, which decreases the intensity from lower depth and affects particularly the intensity based approaches of analysis. Nevertheless, several papers on this topic have been published, e.g. [8].

# Acknowledgments

This work has been supported by European Regional Development Fund - Project FNUSA-ICRC (No.CZ.1.05/1.1.00/02.0123) under Ministry of Education, Youth and Sports.

# References

[1] Lev Dykman and Nikolai Khlebtsov. Gold nanoparticles in biomedical applications: recent advances and perspectives. *Chemical Society reviews*, 41(6):2256–82, March 2012.

[2] Liu Gangjung. *Selected Topics in Optical Coherence Tomography*. InTech, 2012.

[3] Kirill V Larin, Joel D Morrisett, Mohamad G Ghosn, and Valery V Tuchin. Noninvasive functional imaging of tissue abnormalities using optical coherence tomography. In *2010 IEEE Sensors*, pages 2283–2286, November 2010.

[4] Adrian Mariampillai, Beau a Standish, Eduardo H Moriyama, Mamta Khurana, Nigel R Munce, Michael K Leung, James Jiang, Alex Cable, Brian C Wilson, I Alex Vitkin, and Victor X D Yang. Speckle variance detection of microvasculature using swept-source optical coherence tomography. *Optics letters*, 33(13):1530–2, July 2008.

[5] Adrian Mariampillai, Michael K K Leung, Mark Jarvi, Beau A Standish, Kenneth Lee, Brian C Wilson, Alex Vitkin, and Victor X D Yang. Optimized speckle variance OCT imaging of microvasculature. *Optics letters*, 35(8):1257–9, April 2010.

[6] Petra Podlipná and Radim Kolář. Single fiber perfusion phantom for optical coherence tomography. In *European Conferences on Biomedical Optics*, page 88020Q. International Society for Optics and Photonics, June 2013.

[7] Y. Ponce de León, J. L. Pichardo-Molina, N. Alcalá Ochoa, and D. Luna-Moreno. Contrast Enhancement of Optical Coherence Tomography Images Using Branched Gold Nanoparticles. *Journal of Nanomaterials*, pages 1–9, 2012.

[8] K A Vermeer, J Mo, J J A Weda, H G Lemij, and J F de Boer. Depth-resolved model-based reconstruction of attenuation coefficients in optical coherence tomography. *Biomedical optics express*, 5(1):322–37, December 2013.

# Poster Session 2:  Machine Learning, Simulation, and Image Analysis Applications

# Combining EMD with ICA to analyze combined EEG-fMRI Data

Saad M. H. Al-Baddai[1,2]
saad.albaddai@yahoo.com
Karema S. A. Al-Subari[1,2]
s.karema@yahoo.com
Ana Maria Tomé[3]
ana@ua.pt
Gregor Volberg[4]
gregor.volberg@psychologie.uni-regensburg.de
Elmar W. Lang[1]
elmar.lang@ur.de

[1] CIML Group
Institute of Biophysics
University of Regensburg
93040 Regensburg, Germany

[2] Institute of Information Science
University of Regensburg
93040 Regensburg, Germany

[3] IEETA
Dept. de Electrónica e
Telecomunicações
Universidade de Aveiro
3810-Aveiro, Portugal

[4] Institute of Experimental Psychology
University of Regensburg
93040 Regensburg, Germany

**Abstract**

Within a combined EEG-fMRI study of contour integration, we analyze responses to Gabor stimuli with a combined Empirical Mode Decomposition and an Independent Component Analysis. Generaly, responses to different stimuli are very similar thus hard to differentiate. EMD and ICA are used intermingled and not simply in a sequential way. This novel combination helps to suppress redundant modes resulting from an application of ensemble EMD alone. The simulation results show an improved mode separation quality. Hence, the proposed method is an efficient data analysis tool to clearly reveal differences between similar response signals and activity distributions.

## 1 Introduction

It had been already noted by early Gestalt psychologists that the human visual system tends to group local stimulus elements into global wholes. Such grouping is often based on simple rules such as similarity, proximity, or good continuation of the local elements [3]. One special instance of perceptual grouping is contour integration where local parts of Gabor elements are re-integrated into a continuous contour line. Contour integration is typically tested with a stimulus paradigm where arrays of Gabor patches are presented to the subjects (see Fig.1). In 1998, N.E. Huang et al. [4] invented a heuristic tool for complex, non-stationary signal analysis named Empirical Mode Decomposition (EMD). Soon afterwards, EMD has been extended to Ensemble EMD [8], a noise-assisted variant to suppress mode mixing, as well as to multi-dimensional EMD (MEEMD) [6], [9] including complex-valued

©2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

data sets [7]. EEMD/MEEMD cannot guarantee completeness and orthogonality, and may produce redundant components which affect the accuracy of the decompositions. Combining Independent Component Analysis (ICA) with EEMD [5] allows to overcome some of the above limitations. But MEEMD-ICA cannot reveal small differences between closely similar signals. The aim of this paper is to introduce a new way to combine EEMD with ICA which allows to fully suppress redundant signal components in different IMFs. Eventually, the non-stationary signal or image can be decomposed and distinguished accurately, yielding a much better decomposition and feature extraction performances.

# 2 Methods

## 2.1 EEMD/MEEMD

EMD was developed from the assumption that from any signal locally simple oscillations can be extracted. The resulting component signals are called Intrinsic Mode Functions (IMF). Such IMFs are obtained from the signal by means of a sifting algorithm resulting locally in pure oscillations with zero mean. Amplitude and frequency of the IMFs may change over time. Furthermore, IMFs are ordered according to their frequency content. In contrary with wavelets, EMD is a data driven algorithm that decomposes the signal without prior knowledge.

The decomposition of an image, for example an fMRI brain slice, starts by applying EEMD to each column $X_{*n} \equiv \mathbf{x}_n$ of the $M \times N$ - dimensional data matrix $\mathbf{X}$, where $M$ denotes the number of samples and $N$ gives the dimension of the data vectors. The 1D-EEMD decomposition of the $n$-th column becomes

$$\mathbf{x}_n := X_{*,n} = \sum_{j=1}^{J} C_{*,n}^{(j)} \tag{1}$$

where the column vector $C_{*,n}^{(J)}$ represents the residuum of the $n$-th column vector of the data matrix. This finally results in $J$ component matrices, each one containing the $j$-th component of every column $\mathbf{x}_n, n = 1,\ldots,N$ of the data matrix $\mathbf{X}$.

$$\mathbf{C}^{(j)} \quad = \quad [\mathbf{c}_1^{(j)}\ \mathbf{c}_2^{(j)}\ \cdots\ \mathbf{c}_N^{(j)}] = [C_{*,1}^{(j)}\ C_{*,2}^{(j)}\ \cdots\ C_{*,N}^{(j)}] \tag{2}$$

Next one applies an EEMD to each row of eqn. 2 yielding

$$C_{m,*}^{(j)} \quad = \quad \left(c_{m,1}^{(j)} c_{m,2}^{(j)} \cdots c_{m,N}^{(j)}\right) = \sum_{k=1}^{K} \left(h_{m,1}^{(j,k)} h_{m,2}^{(j,k)} \cdots h_{m,N}^{(j,k)}\right) = \sum_{k=1}^{K} H_{m,*}^{(j,k)} \tag{3}$$

where $c_{m,n}^{(j)} = \sum_{k=1}^{K} h_{m,n}^{(j,k)}$ represents the decomposition of the rows of matrix $\mathbf{C}^{(j)}$. These components $h_{m,n}^{(j,k)}$ can be arranged into a matrix $\mathbf{H}^{(j,k)}$ according to

$$\mathbf{H}^{(j,k)} = \begin{bmatrix} h_{1,1}^{(j,k)} & h_{1,2}^{(j,k)} & \cdots & h_{1,N}^{(j,k)} \\ h_{2,1}^{(j,k)} & h_{2,2}^{(j,k)} & \cdots & h_{2,N}^{(j,k)} \\ \vdots & \vdots & \cdots & \vdots \\ h_{M,1}^{(j,k)} & h_{M,2}^{(j,k)} & \cdots & h_{M,N}^{(j,k)} \end{bmatrix} \tag{4}$$

The resulting component matrices have to be summed to obtain

$$\mathbf{C}^{(j)} = \sum_{k=1}^{K} \mathbf{H}^{(j,k)}. \tag{5}$$

Finally this yields the following decomposition of the original data matrix $\mathbf{X}$

$$\mathbf{X} = \sum_{j=1}^{J} \mathbf{C}^{(j)} = \sum_{j=1}^{J} \sum_{k=1}^{K} \mathbf{H}^{(j,k)} \tag{6}$$

where each element is given by

$$x_{m,n} = \sum_{j=1}^{J} \sum_{k=1}^{K} h_{m,n}^{(j,k)} \tag{7}$$

To yield meaningful results, components $h_{m,n}^{(j,k)}$ with comparable scales, i. e. similar spatial frequencies of their textures, should finally be combined [9] according to comparable minimal scale combination principle(CMSC). In practice, for two-dimensional data sets this implies that the components of each row, which represent a common horizontal scale, and the components of each column, which represent a common vertical scale, should be summed up [9].

Hence, the CMSC - principle leads to BIMFs given by

$$\mathbf{S}^{(k')} = \sum_{k=1}^{K} \mathbf{H}^{(k,k')} + \sum_{j=k'+1}^{J} \mathbf{H}^{(k',j)} \tag{8}$$

which thus yields a decomposition of the original data matrix $\mathbf{X}$ into BIMFs according to

$$\mathbf{X} = \sum_{k'=1}^{K} \mathbf{S}^{(k')} \tag{9}$$

where $\mathbf{S}^{(K)}$ represents the non-oscillating residuum. The extracted BIMFs can be considered features of the data set which, according to the CMSC - principle, reveal local textures with characteristic spatial frequencies which help to discriminate the functional images under study.

## 2.2 ICA

Independent component analysis(ICA) aims to find a linear representation of data based on maximally non-Gassian components which renders them statistically independent. Consider $M$ statistically independent sources $\mathbf{H} = \{\mathbf{h}_1, \mathbf{h}_2, ...., \mathbf{h}_M\}$ and $M$ sensor signals $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, ...., \mathbf{x}_M\}$. The goal of ICA is to find a de-mixing matrix which recovers the hidden components $H$ underlying the observations $X$ . The different $\mathbf{y}_m$ are estimates of the latent variables $\mathbf{h}_m$ according to the following model

$$\mathbf{y}_m = \mathbf{W}\mathbf{x}_m = \mathbf{W}\mathbf{A}\mathbf{h}_m \qquad \text{with } \mathbf{WA} = \mathbf{PD}. \tag{10}$$

where $W$ denotes the demixing matrix, $P$ represents a permutation matrix and $D$ a scaling matrix. Examples of common ICA implementations are the *JADE* algorithm [2] and the *INFOMAX* algorithm [1].

## 2.3 Proposed Method

As mentioned above, a combination of EEMD/MEEMD with ICA yields a signal decomposition free of redundant remnants of other components in any extracted IMF. The newly proposed methods works as follows:

1. Decompose the measured signal or image $\mathbf{X}$ with EEMD and BEEMD resulting in IMFs or BIMFs which are ordered according to their frequency content.

2. Initialize $i = 1$ and $j = 1$; where $i, j = 1.....K$ and $K$ represents the number of extracted modes.
3. Choose a pair of IMFs, $IMF_i$ and $IMF_{j\neq i}$, and feed it into an ICA algorithm. After decomposing the IMFs with ICA, we obtain two independent components $IC_i$ and $IC_j$, respectively.
4. If $i > j$, choose the IC with higher frequency according to their Hilbert - Huang Transform [4]. Otherwise, choose the IC with the lower frequency.
5. Replace $IMF_i$ with the selected IC.
6. Increase $j$ by one and repeat the steps above until $j = K$. This results in a new Intrinsic Mode Component (IMC).
7. Increase $i$ by one and repeat steps 3 to 6 until $i = K$.
8. This procedure yields IMCs/BIMCs which neither fulfill the conditions for an IMF or a BIMF nor an IC.

# 3 Simulation

In order to demonstrate the performance of the proposed method, one EEG signal and a slice of the related fMRI image collected during a contour integration task is selected. Such signals are illustrated for both stimuli, *contour* and *non-contour* (see Fig. 1). The EEG signals and related fMRI images, as shown in Fig. 1, are decomposed by EEMD in case of EEG signals, and by BEEMD in case of fMRI images. Eight IMFs are extracted from the EEG signals (but only three are shown), and six IMFs are obtained from each fMRI image. The resulting components are shown in Fig. 2. As can be seen from the original signals, exhibited in Fig.1, no noticeable differences between the recorded signals, following contour and non-contour stimuli, can be detected. Even from the IMFs and BIMFs, obtained from an EEMD/BEEMD decomposition (shown in the top row of Fig.2) no characteristic difference can be noticed. A similar result is obtained if, after decomposing the recorded signals with EEMD/BEEMD, an ICA is applied to the IMFs/BIMFs directly (see Fig. 2, middle row). This is due to incomplete signal decompositions by these methods which leave remnants of one components in others, causing some redundancy in the different components. Such redundancies load a large subjectivity onto any diagnosis based on these methods. In contrast, the IMCs/BIMCs extracted by our proposed method, can overcome this limitation and yields clearly different characteristics corresponding to both stimuli. This can be seen clearly from the bottom row of Fig. 2. As a cross-check, we can obtain ICs identical to the ones shown in Fig. 2, middle row, if we apply ICA to the IMCs/BIMCs obtained from our method. This corroborates that no information loss has occurred during the analysis.

# 4 Discussion and Conclusion

In real applications, stimulus responses would most likely not be identical under different conditions. Rather some response asymmetries are to be expected. Because response differences are small, they become submerged in the background of the extracted modes. Hence, such differences cannot be classified simply by visual inspection of the responses. We even demonstrated that such small differences cannot be revealed by plain EEMD/BEEMD or a sequential combination of EEMD with ICA. The reason is some partial mode mixing which appears in noise-assisted ensemble EMD. Our proposed method is able to do so because it helps to suppress remnants of other modes interfering in any of the extracted modes. This

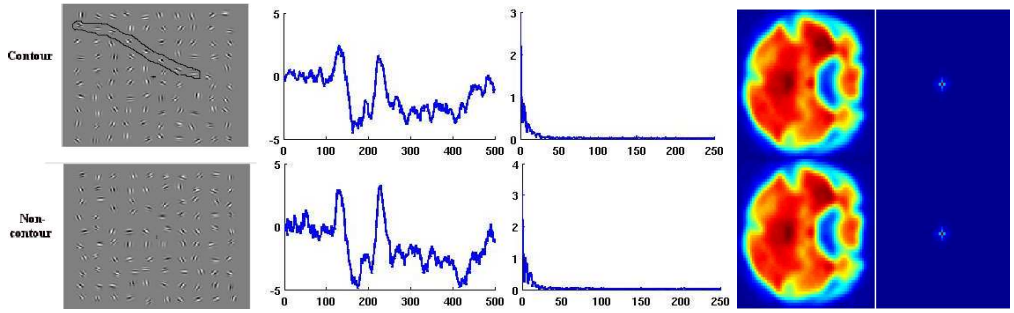



Figure 1: *Left:* contour and non-contour stimuli. *Middle:* EEG signals and their corresponding Fourier spectra. *Right:* stimulus-related fMRI images and their spatial frequency spectra.

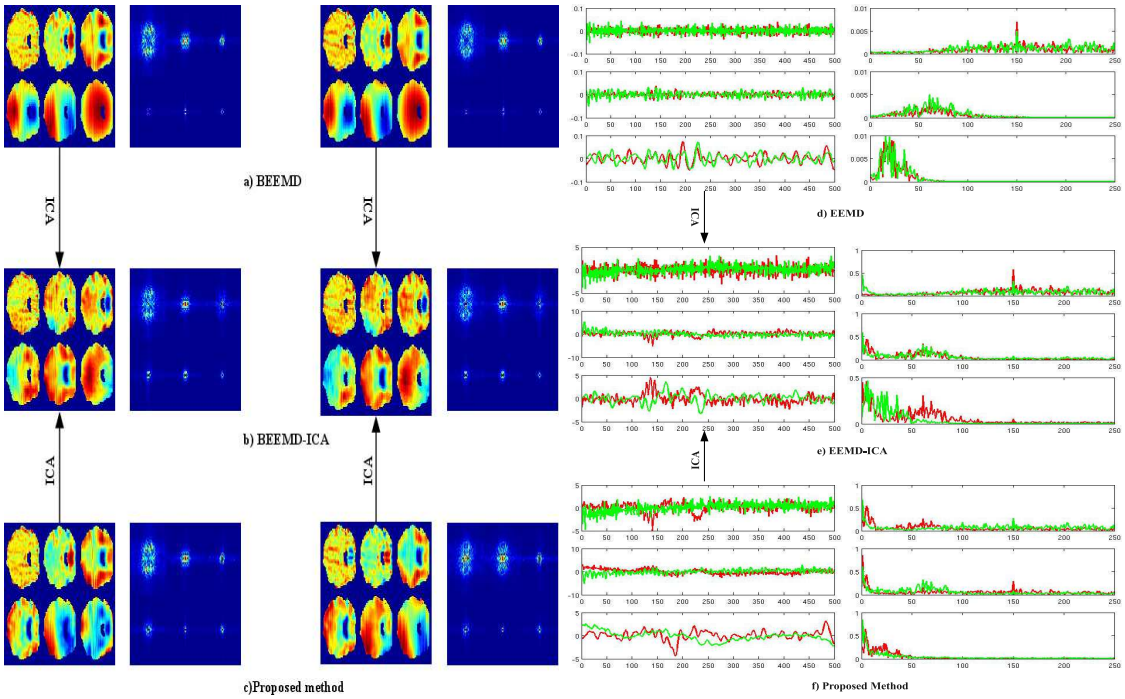



Figure 2: fMRI activity distributions and EEG recordings in response to contour (column 1 and column 3, red line) and non-contour (column 2 and column 3, green line) stimuli. *Top:* BIMFs and related IMFs extracted with BEEMD and EEMD from fMRI and EEG recordings. *Middle:* ICs resulting from an ICA applied to BIMFs and related IMFs obtained from original data sets directly. *Bottom:* BIMCs and related IMCs extracted with our proposed method. For fMRI images, modes are sorted from left to right and from top to bottom according to their spatial frequency content. For EEG time series, the three interesting modes are shown together with their corresponding Fourier spectra.

results in clean modes with no interferences from other modes and thus improves the separation quality considerably. The results showed that the proposed method can be applied to efficiently extract features from biomedical signals and images. This is especially important if different response classes need to be differentiated.

Response classification of our proposed method has been evaluated against competing methods like applying BEEMD to the raw data sets or applying BEEMD and ICA sequentially. The study comprised 18 subjects where a combined EEG-fMRI analysis has been performed within a contour integration task with contour and non-contour Gabor stimuli. As classifier, a Support Vector Machine (SVM) using the Leave One Out Cross Validation (LOOCV) technique has been employed . Dimension reduction has been achieved by projecting the extracted modes onto principal components and using the projections as input to the classifier. A Student t-test was used to select informative features and the parameters of the classifier were optimized by using a grid search approach. The values between square brackets in Table 1 show the number of selected features for an optimal performance and the first column indicates the mode.

Table 1: Comparison of statistical measures (Accuracy, Specificity and Sensitivity) obtained with different techniques evaluating corresponding classification results.

| | BEEMD | | | BEEMD-ICA | | | Proposed Method | | |
|---|---|---|---|---|---|---|---|---|---|
| ♯ | Acc | Spec | Sens | Acc | Spec | Sens | Acc | Spec | Sens |
| 1 | 0.81[7] | 0.84 | 0.79 | **0.84**[**34**] | **0.80** | **0.89** | **0.92**[**35**] | **0.89** | **0.94** |
| 2 | 0.82[4] | 0.84 | 0.79 | 0.68[1,30] | 0.68 | 0.68 | 0.63[1] | 0.58 | 0.68 |
| 3 | **0.89**[**2**,**11**] | **0.89** | **0.89** | 0.79[23] | 0.79 | 0.79 | 0.71[32] | 0.78 | 0.63 |
| 4 | 0.84[3] | 0.84 | 0.84 | 0.63[22] | 0.68 | 0.57 | 0.84[3] | 0.84 | 0.84 |
| 5 | 0.84[29] | 0.89 | 0.79 | 0.66[18] | 0.63 | 0.68 | 0.74[29] | 0.74 | 0.74 |
| 6 | 0.79[26] | 0.74 | 0.84 | 0.76[29] | 0.84 | 0.69 | 0.71[21] | 0.68 | 0.74 |

# References


[1] T. Bell and T. Seinoswki. An information-maximization approach to blind source separation and blind deconvolution. *Neural Computation*, 7(6):1004 – 1034, 1995.

[2] J.F. Cardoso, France Telecom Paris, and A. Souloumiac. Blind beamforming for non-gaussian signals. *Radar and Signal Processing , IEE Proceedings F*, 140(6):362–370, 1993.

[3] Donderi DC. Visual complexity: a review. *Psych Bull*, 132:73–97, 2006.

[4] N. E. Huang, Z. Shen, S. R. Long, M. L. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. Roy. Soc. London A*, 454:903–995, 1998.

[5] B. Mijović, M.De Vos, I. Gligorijević, and J. Taelman. Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis. *IEEE Transactions on Biomedical Engineering*, 57(9):2188–2196, 2010.

[6] J.C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang, and Ph Bunel. Image analysis by bidimensional empirical mode decomposition. *Image Vis. Comput.*, 21(12), 2003.

[7] T. Tanaka and D. P. Mandic. Complex empirical mode decomposition. *IEEE Signal Processing Letters*, 14(2): 101–104, 2006.

[8] Zh. Wu and N. E. Huang. Ensemble Empirical Mode Decomposition: a noise-assisted data analysis method. *Adv. Adaptive Data Analysis*, 1(1):1–41, 2009.

[9] Zh. Wu, N. E. Huang, and X. Chen. The Multidimensional Ensemble Empirical Mode Decomposition Method. *Adv. Adaptive Data Analysis*, 1:339–372, 2009.

# Automated histological quantification of trabecular bone tissue in critical illness

Thomas Janssens[1]
thomas.janssens2@med.kuleuven.be
Ine Vanhees[1]
ineke.vanhees@med.kuleuven.be
Jan Gunst[1]
jan.gunst@med.kuleuven.be
Helen Owen[2]
h.owen@qmul.ac.uk
Greet Van den Berghe[1]
greet.vanden.berghe@kuleuven.be
Fabian Guiza Grandas[1]
fabian.guiza@med.kuleuven.be

[1] Laboratory of Intensive Care Medicine,
KU Leuven
UZ Herestraat 49, box 7003
3000 Leuven
Belgium

[2] Translational Medicine and
Therapeutics
William Harvey Research Institute
Barts and the London School of
Medicine and Dentistry
Charterhouse Square
London EC1M 6BQ

**Abstract**

Interpretation of histological images is often subjective and time-consuming. Automated approaches offer efficiency and objectivity along with increased precision in results. We focus on the automated quantification of trabecular bone histology images, which are used to study effects of critical illness in a rabbit model. The amount and ratio's of immature and mature bone tissue is of particular interest. We quantify these different tissues through a per-pixel classification model using HSV values. These results are then smoothed with a median filtering step and refined through spatial constraints encoding domain knowledge in order to arrive at a well-performing final segmentation. This allows precise staining areas to be calculated, and individual stain properties to be measured.

## 1 Introduction

Over the past decades, improvements in intensive care medicine has allowed a growing number of patients to survive previously lethal insults such as surgery, trauma or burn injury. However, these advances also result in a larger population of patients that are in a state of prolonged critical illness, remaining dependent on vital organ support for weeks or months [5, 7].

Prolonged critically ill patients also have distinct alterations in their bone metabolism, which can result in pronounced bone loss, impaired traumatic or surgical fracture healing, and osteoporosis [8]. Our group studied bone degradation in a previously developed rabbit model for critical illness. Hundreds of Masson Trichrome-stained images were obtained from 15 critically ill and control rabbits at 10x magnification. These images show sections containing both the immature bone (dark red) and demineralized, mature bone (blue); some

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

examples are shown in figure 1. The relative presence of these types of tissue needs to be accurately quantified to properly assess the effects of critical illness on bone metabolism [9].

However, manually evaluating these images poses known difficulties. Firstly, the amount generated is often too large for an in-depth evaluation, so in practice all images are often evaluated on a very rough scale (e.g. a discrete rating of 0-3 for presence of a tissue type). Even so, evaluating these images is very time consuming. Secondly, it is very hard for human evaluators to give an objective judgment due to the many different structures and color levels present. Often, there exists some variation in interpretation given by different evaluators, or by one evaluator over a period of time. One of the main problems is distinguishing immature bone from bone marrow. The former is dark red and adjacent to blue-stained mature bone, while the latter is present across the entire image and has a stained color ranging from dark to bright red.

There exist some semi-automated approaches, such as measuring the fraction of image pixels that crosses a threshold level on one of the images color channels (i.e. detecting a red staining by quantifying the amount of pixels above a certain threshold on the red channel). These methods might accelerate the process slightly, but still require a significant amount of time and contain a degree of subjectivity.

Automatically processing these images is non-trivial because of the large variation in morphology presented by the bone tissue. The structures are densely present in the image, unpredictably positioned, shaped and sized, with variable staining intensity and imperfect separation between the tissues' stain colors.

We have opted for a machine learning method as they are known to be well suited to detecting the underlying properties of large datasets (visual or otherwise). This knowledge can then be applied in classifying new, unseen examples. Such methods have been shown to have good results in the histological domain [4, 6].

We describe a methodology to segment these images in a fully automatic way by combining pixel-level classification through a machine learning model with larger scale morphological domain knowledge. We compare this to existing semi-automated methods for trabecular bone quantification such as described above and with manual scoring by domain experts.

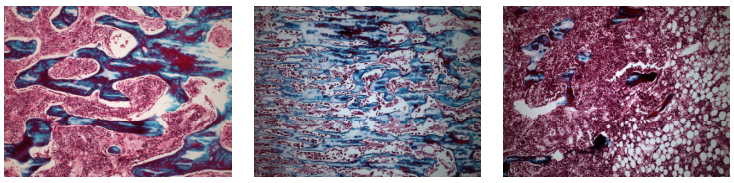


Figure 1: Some examples of histological bone images.

## 2 Methodology

Our framework consists of two parts: a pixel-wise classification step and morphological post-processing; we discuss each step in turn below. This approach can be contrasted with the semi-automatic one, where the user splits the image into its red, blue and green channels, and defines an image-specific threshold value for each channel, in order to extract the red and blue stained regions. In such an approach no steps are taken to avoid the inclusion of similarly-stained bone marrow in the immature bone area count, often resulting in an

overestimation of the latter's area. The two main advantages our method offers are the full automation, and the capacity to take the above mentioned tissue difference into account.

## 2.1 Pixel-wise Classification

A suitably representative image is selected from the dataset and used to train the classifier. We assign all its pixels a class label $C = \{Red, Blue, Other\}$. We choose these color-based categories over tissue type because it is not yet possible at the single pixel stage to make a full distinction between immature bone and marrow. A random forest classier [2] is then trained on a randomly selected subsample of $n = 10^5$ pixels, as larger subsets were not found to introduce any meaningful improvement in model performance.

In this model, each pixel is characterized by its *Hue*, *Saturation* and *Value* components. Though this is a very limited feature set, the high level of class information contained in a pixel's color values results in a classification accuracy of 94.3%, measured by 10-fold cross-validation. Though this evaluation was only performed on one image, the results extrapolate well to others in the dataset, as we discuss in sections 3 and 4.

## 2.2 Post-processing

After the initial pixel classification a number of post-processing steps are applied to improve the accuracy of the result. The pixel-wise nature of the classification leads to many isolated pixel classifications, where the majority of its neighbors have a different class. These single pixels can be assumed to be noise, since the desired cellular structures are presumed to have an area on a scale of hundreds of pixels. Therefore we apply a median filtering step, giving each pixel the class of the majority of its neighbors in an $(2f+1)$-by-$(2f+1)$ square around it, where $f$ typically is 1 or 2. This way, each pixel is assigned the most common class value present it the set of itself and its 8-connected neighbors.

After this step, we impose a minimum and maximum size for identified bone tissue structures. On the advice of the domain expert, we disregard all segments identified as immature and mature bone under size 50px and 300px, respectively.

Finally we impose spatial constraints by only recognizing *Red* segments as being immature bone if they are directly adjacent to (i.e. have a shared border) with a remaining *Blue* segment. This is necessary as it eliminates the visually very similar bone marrow segments, that were also identified as belonging to the *Red* class.

After this post processing, we obtain our resulting segmented image. These steps are summarized in figure 2.

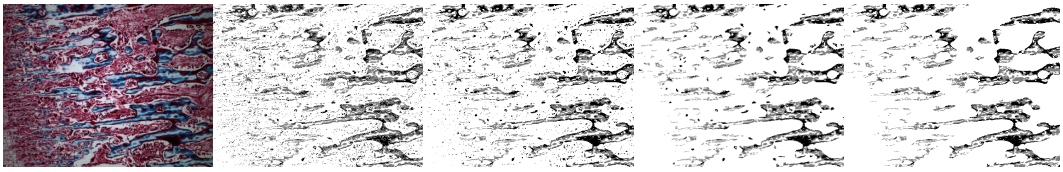
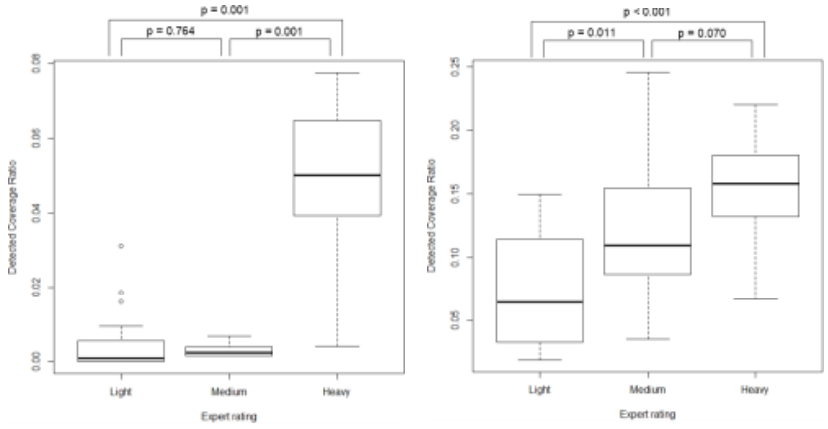


Figure 2: Illustration of the different parts of our method: Unprocessed, after pixel classification, after median filtering, after eliminating small red and blue segments, and finally after elimination based on spatial constraints.

# 3 Application

Our framework has been applied to the quantification of 341 trabecular bone images taken from 15 critically ill rabbits. These processed images were evaluated and approved by domain experts. Sections from the right proximal tibia from 15 healthy controls and prolonged critically ill rabbits were stained by Masson's Trichrome stain in order to analyze the amount of demineralized, mature bone vs. unmineralized osteoid or immature bone. These slides were then imaged using fluorescence microscopy.

Of these images, 36 were also scored visually by three experts from our lab for presence of immature and mature bone tissue, rated in discrete categories {*light*, *medium*, *heavy*}. Figure 3 plots the results of our analysis against these scores to give an indication of the agreement between these discrete expert evaluations and the continuous values derived with our method. As the figure shows, classes are differentiated well (with a significance level of $p = 0.05$).




(a) Immature bone (b) Mature bone

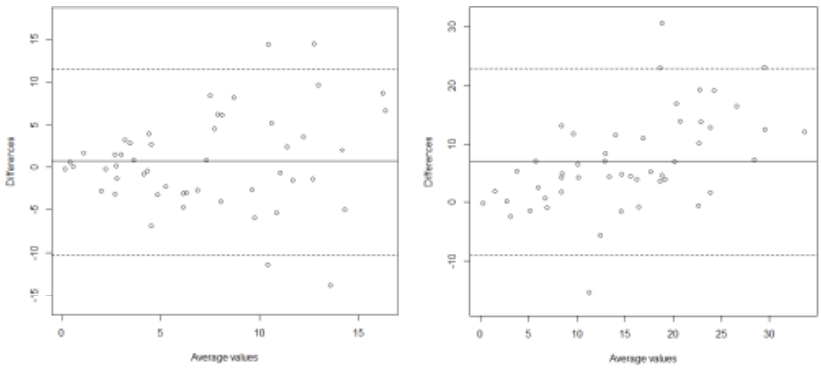
Figure 3: Human classification (median of three) of 36 trabecular bone images vs. relative image coverage area as quantified by our method.

We also need to note that there exists a non-negligible variability in opinion between these experts. We measure the inter-observer agreement by using Fleiss' Kappa statistic [3]. A value of $\kappa = 0$ indicates no better agreement that would happen by random chance, while $\kappa = 1$ represents perfect agreement. For our three raters the quantification of the presence of mature (blue) and immature(red) bone, rated with the three categories mentioned above, is $\kappa_{Mature} = 0.795$ (95 % CI = [0.760,0.831]) and $\kappa_{Immature} = 0.289$ (95 % CI = [0.251, 0.314]).

To compare our method to the more time-intensive manual threshold-based approach, where the user selects appropriate thresholds on the image color channels, we use Bland-Altman plots [1], shown in figure 4 for 50 images.

# 4 Discussion

We analyzed 341 images for the presence of mature and immature bone and compared this to the evaluation given by experts with both a qualitative and semi-automated quantitative method.




(a) Immature bone (b) Mature bone

Figure 4: Bland-Altman plots for (a) immature (b) and mature bone. The horizontal axis represents percentage of image area covered (mean of both measurements), while the vertical axis shows the difference in measurement values (manual - automatic).

The results of this method were approved by domain experts and used for quantification of bone tissue from a rabbit model [9], as they were determined to be more accurate than both other methods.

This is confirmed by the plots in figure 4. They show that the manual threshold-based method generally gives slight overestimation of immature bone tissue area are also slightly higher (0.68% more, figure 4(a)) and they diverge in proportion to the image area covered. It also gives a higher estimate for the mature bone tissue area (6.95% more, figure 4(b)). However, this method was designed as an ad-hoc solution, and shouldn't be considered a gold standard. As simple thresholding only takes into account pixel color and no structural information, in some cases it can be impossible not to over- or underestimate the areas involved. Since our results were visually confirmed to be more accurate by the domain experts, these plots can be seen as evidence of the limited accuracy of even this more quantitative manual method.

If we compare our method to the more granular categorical division made, we notice a fair correspondence with the given ordinal labels (see figure 3). Though this distinction is worse for the light and medium density immature bone, the low Fleiss' kappa value there indicates disagreement among the different raters as well.

Some user interaction is required in setting up the pipeline, as the post-processing steps have to conform to the domain requirements, and the user has to be aware of their operation and possibilities. However, once this pipeline is set up the algorithm can robustly classify large batches of images. By using relatively simple features as we did in our experiment, hundreds of images can be processed in a matter of minutes.

Despite only being trained and evaluated on one image, the robustness of the model is demonstrated by its good performance on a wide variety of images with unequal staining intensity and variable lighting. Though perhaps a wider variety in training images might increase the performance even further, the fact that an expert only need label one image to get usable results speaks to the user-friendliness of the algorithm.

To refine the results of our method, a logical next step would be to introduce some new features and perhaps redefine the classes to map onto the tissue types instead of color. Textural or spatial information would also be beneficial.

# 5 Conclusions and Future Work

We have developed a extensible framework capable of using information at the pixel scale and morphological domain knowledge for detecting specific regions of interest in Masson Trichrome-stained trabecular bone histological images. The tissue areas detected matched expert opinion and were used to evaluate a rabbit model experiment [9].

We envision the possibility of this processing pipeline being of use in quantifying other images with large variation in shape and size of regions of interest. The pixel-wise classification is independent of overall region shape of size, while the morphological rules after this step do allow selections of regions based on their spatial properties and relation to other regions.

# References

[1] Martin J Bland and Douglas G Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327(8476):307–310, 1986.

[2] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[3] Joseph L Fleiss. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378, 1971.

[4] Lei He, L Rodney Long, Sameer Antani, and G Thoma. Computer assisted diagnosis in histopathology. *Sequence and Genome Analysis: Methods and Applications*, pages 271–287, 2010.

[5] Greet Hermans, Ilse Vanhorebeek, Sarah Derde, and Greet Van den Berghe. Metabolic aspects of critical illness polyneuromyopathy. *Critical care medicine*, 37(10):S391–S397, 2009.

[6] Nikita Orlov, Josiah Johnston, Tomasz Macura, Lior Shamir, and Ilya Goldberg. Computer vision for microscopy applications. *Vision Systems: Segmentation and Pattern Recognition*, pages 221–242, 2007.

[7] Greet Van den Berghe, Francis de Zegher, and Roger Bouillon. Acute and prolonged critical illness as different neuroendocrine paradigms 1. *Journal of Clinical Endocrinology & Metabolism*, 83(6):1827–1834, 1998.

[8] Greet Van den Berghe, David Van Roosbroeck, Philippe Vanhove, Pieter J Wouters, Lutgart De Pourcq, and Roger Bouillon. Bone turnover in prolonged critical illness: effect of vitamin d. *Journal of Clinical Endocrinology & Metabolism*, 88(10):4623–4632, 2003.

[9] Ine Vanhees, Jan Gunst, Thomas Janssens, Andy Wauters, Erik Van Herck, Sophie Van Cromphaut, Greet Van den Berghe, and Helen C Owen. Enhanced immunoreceptor tyrosine-based activation motif signaling is related to pathological bone resorption during critical illness. *Hormone and Metabolic Research*, 63(12):862–869, 2013.

# Analysis of capillary-like structures formed by endothelial cells in a novel organotypic assay developed from heart tissue

Simon Bowyer [1]
simonbowyer@gmail.com
Chryso Kanthou [2]
c.kanthou@sheffield.ac.uk
C. C. Reyes-Aldasoro [3]
reyes@city.ac.uk
http://staff.city.ac.uk/~sbbk034/

[1] University of Sussex, Brighton UK
[2] Tumour Microcirculation Group, The University of Sheffield, Sheffield, UK
[3] Biomedical Engineering Research Group, City University London, London UK

**Abstract**

In this work, we present an algorithm to perform automatic segmentation, feature extraction and dosage classification of images derived from a novel organotypic angiogenesis assay developed to assess the effects of ionising radiation in the mouse heart. The images presented very different conditions of illumination and conditions of density and shapes of cells. The algorithm consists of a pipeline of several steps, which were validated against hand-segmented images. The algorithm provided satisfactory results, as all images were correctly dose-classified. The cells exposed to the lowest radiation dose were observed to have the greatest relative feature variability.

# 1 Introduction

Formation of new blood vessels by angiogenesis is a feature of embryonic development, while in the adult, angiogenesis is associated with several pathological conditions such as cancer, as well as normal physiological tissue repair processes. For example, in the heart, injury triggers repair mechanisms within the myocardium, which are dependent on angiogenesis. For example, following a myocardial infarction, ischaemia triggers vascular responses and new blood vessel formation/angiogenesis aiming to repair the damage and re-establish flow [1]. Various organotypic assays have been developed in order to model the angiogenic process, test therapeutic agents and analyse the molecular mechanisms involved [2]. In these assays, which are generally based on co-culturing endothelial cells with fibroblasts, the endothelial cells migrate and align to form capillary-like structures embedded within fibroblasts and matrix. We have recently developed a heart angiogenesis organotypic assay by culturing enzymatically digested heart tissue fragments [3]. In this

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

model, cardiac endothelial cells form tubule-like structures embedded within the various cellular components of the digested heart tissue, which include fibroblasts and pericytes. We used this model system to establish the effects of ionising radiation on heart microvascular damage and inhibition of angiogenesis. This paper is concerned with the geometric analysis of these tube-like structures using cultures derived from the hearts of C67BL/6 mice that were irradiated with a range of ionizing radiation doses (0.2, 2, 8, 16 Gy) [n=4x4] and sham-irradiated controls [n=4]. The hearts were excised for assay at 20 weeks post irradiation. One of the 8 Gy images was discarded due to duplication, leaving 19 images in total.

# 2 Materials and Methods

Organotypic cultures were stained with lectin to identify the endothelial cells and a representative image at each irradiation dose is shown in Fig. 1. The images are greyscale, with low contrast, each 696x520 pixels in size and comprising a cluttered background overlaid with the tube-like structures formed by endothelial cells (white arrow). There are significant inter-image and intra-image shading variations. Dead cells or debris were also present, sometimes occluding the tubules (black arrow). The tubules were the primary target of the analysis, though the debris (biological in nature rather than imaging artefacts) is also relevant.

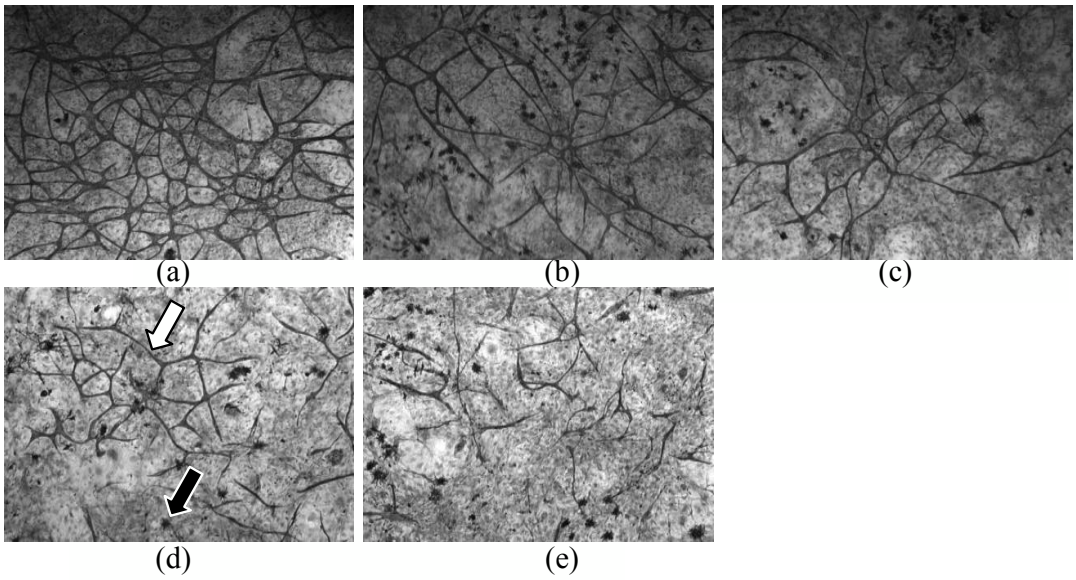
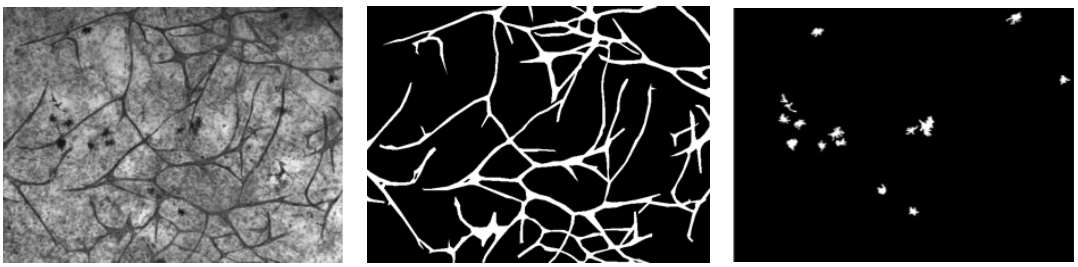



Figure 1: Representative images of the image set; white arrow marks tubule, black arrow marks debris. Irradiation doses: (a) 0 Gy, (b) 0.2 Gy, (c) 2 Gy, (d) 8 Gy, (e) 16 Gy.

The automatic segmentation was validated against manual segmentation (Fig. 2), performed within the MATLAB *roipoly* function by one of the authors without knowledge of the subsequent algorithm or of the automatic segmentation results.



(a) (b) (c)

Figure 2: Manual segmentation (a) original image, (b) tubules, (c) debris.

An automatic segmentation script was developed in MATLAB with the following steps: 1) Median filtering with a 5×5 pixel neighbourhood using *medfilt2*(); 2) Thresholding based on the local mean of each 151×151 neighbourhood, for each pixel > 0.75 $\overline{x}$ and > 1 σ [4, p. 527]; 3) Removal of noise, i.e. objects with area < 25 pixels; and 4) Removal of likely debris, i.e. objects of solidity > 0.6. 'Solidity' is the proportion of the number of pixels in an object's convex hull which are also in that object [5]. This step removes objects that are insufficiently spindly to be tubules. These steps are illustrated in Table 1.

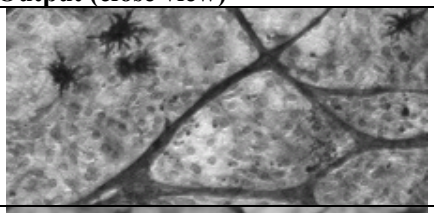
| **MATLAB function** | **Implementation** | **Output (close view)** |
|---|---|---|
| | Original image |  |
| *medfilt2()* | De-noise while retaining edges by applying a median filter with a 5x5 pixel neighbourhood | |
| *localthresh()* | Threshold using local means with 151x151 neighbourhood, std dev multiplier of 1, mean multiplier of 0.75. These parameters produced peak mean segmentation correlation for all images. | |
| *bwarea()* | De-speckle by deleting all binary objects of area < 25 pixels | |
| *regionprops()* | Remove debris, i.e. binary objects with solidity > 0.6 (pixels in object / pixels in convex hull) | |

Table 1: Automatic segmentation algorithm pipeline

The following feature metrics were extracted from the segmented images: 1) Relative tubule area, as a proportion of total image area; 2) Euler number (i.e. number of 'holes' in the tubule mesh) as a proxy for the count of regions completely enclosed by tubules (*bweuler*); 3) Largest major axis of the tubule objects in the image, as a proportion of the diagonal image size; and 4) Largest minor axis of the tubule objects in the image, as a proportion of the diagonal image size. These features were selected for their discrimination properties. No formal means of feature selection was used; we leave that as further work.

The extracted features were classified with the function *treebagger* [6][7], to grow regression and classification decision trees with the intent of automatically placing the target images into dosage categories. The same function was also reapplied with the intent of automatically estimating the dosage exposure as a continuous value.

# 3 Results

Validation was performed as a binary pixels-by-pixel comparison against the manual ground truth (Fig. 3). The results of the classification algorithm showed a mean misclassification for all target images of 0.0733 ±0.0455. Misclassification overall ranged between 0.0263 and 0.2059. The visible output of the automated segmentation method when applied to two sample images is shown in Fig. 4.

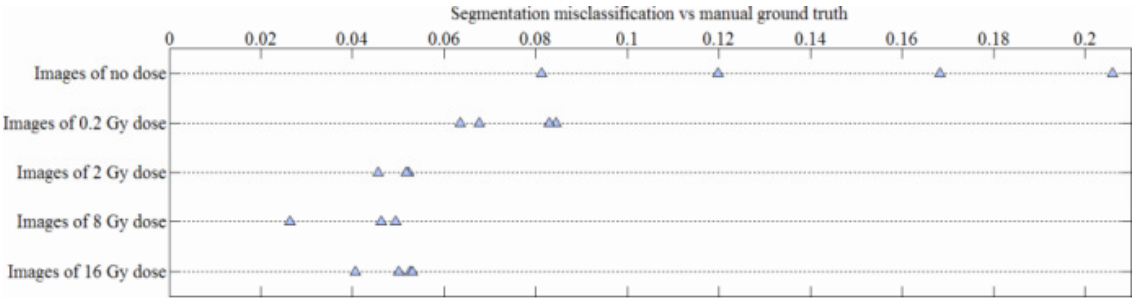



Figure 3: Automatic segmentation accuracy against manual ground truth, grouped by dose. Performance on irradiated images is considerably better than on control images.

Comparison of two of the extracted feature metrics is presented in Fig. 5a (manual segmentation) and Fig. 5b (automatic segmentation). From inspecting Fig. 5a, it is apparent that the largest relative biological morphological variation is present in a healthy specimen exposed to the smallest dose of radiation (0.2 Gy). Even with only two feature dimensions plotted, the topological changes between 0 and 0.2 Gy are greater than those changes that occur between 0.2 and 2 Gy.

The grouping of vector sets into obvious clusters indicates suitability for the application of automatic classification techniques, though 8 Gy and 16 Gy images are difficult to separate when only these two features are examined. It can be seen that the cluster separation is reasonably clear in the data extracted from the manually segmented images (Fig. 5a), and that the clusters are noticeably less separated in the data extracted from the automatically segmented images (Fig. 5b). This indicates that there is low-dimensional information loss in the automatically segmented images due to segmentation inaccuracy.

All automatically segmented images were correctly dose-classified (Table II). The dose-classification confusion matrix of the *treebagger* was extremely good; this is thought to be

due to the dose-discrimination exhibited by the extracted features. Though *treebagger* (being a random forest algorithm) assigns ⅔ of inputs to training and ⅓ to testing and so cross-validation is not required, there were few images and no measurement of misclassification cost was attempted. The feature extraction appeared to be robust to the misclassifications of the segmentation process. Mean classification confidence is 0.83± 0.11. Minimum classification confidence is 0.66, for one image of 2 Gy dosage.

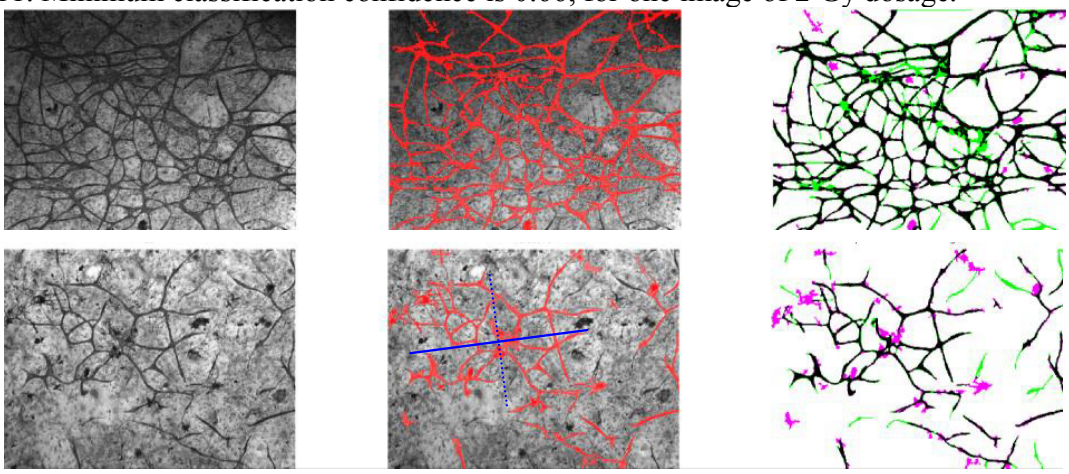


Figure 4: Segmentation results: (left) representative images; (centre) automatically-segmented tubules overlaid in red; (right), validation: correct segmentation (black) false positives (magenta), false negatives (green). Largest second-moment major and minor axis region features are plotted as solid and dotted blue respectively (centre lower), these features are thought to assist with measurement of tubule fragmentation. A dark image seems to have missed tubules, whilst a bright image segmented background and debris as tubules.

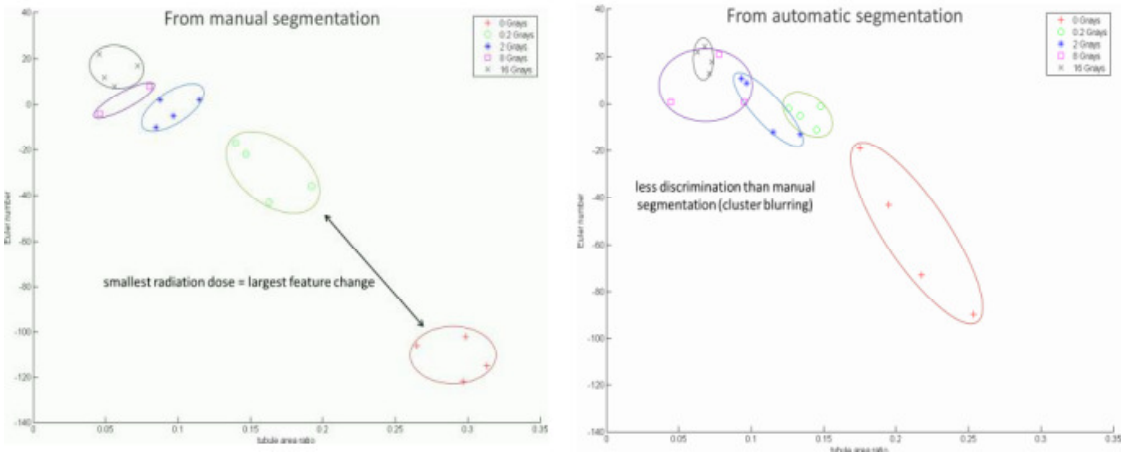



Figure 5: (a) Plot of Euler number vs. tubule area ratio extracted from manual segmentations. There is an inverse relationship between tubule area ratio and Euler number. Note the cluster discrimination, and that the largest feature change is caused by the smallest radiation dose. (b) Plot of Euler number vs. tubule area ratio extracted from automatic segmentations. Note the reduced cluster discrimination in comparison to the manual segmentation in Fig 5(a).

| | Classification category | | | | |
|---|---|---|---|---|---|
| **Images at dose, Gy** | **'0'** | **'0.2'** | **'2'** | **'8'** | **'16'** |
| **0** | **0.88** | 0.08 | 0.04 | 0.02 | |
| **0.2** | 0.05 | **0.78** | 0.14 | 0.06 | |
| **2** | 0.07 | 0.11 | **0.76** | 0.21 | 0.02 |
| **8** | 0.01 | 0.05 | 0.15 | **0.77** | 0.12 |
| **16** | | | 0.02 | 0.04 | **0.96** |

Table 2: Dose-classification mean confusion matrix derived from automatic segmentation

# 4 Conclusion

Automated segmentation of the source images was mildly successful, with an automated algorithm pipeline giving pixel-level misclassification versus manually segmented images of minimum 0.026, mean 0.073, and maximum 0.206. Four spatial features with high dosage discrimination were identified and then applied to reduce the segmented images to low-dimensional form. The smallest radiation dosages cause the largest relative cell structure feature changes. This is clearest when extracting features based on manual segmentation, though is still apparent when using automated segmentation. All images were correctly automatically dose-classified with > 65% confidence and five images with > 95% confidence. This indicates that an automatic segmentation and classification algorithm can be applied to classify and investigate the characteristics of the effects of radiation on cardiac tissue. We plan to extend the algorithm, investigating the effect of the shading [8] on the segmentation and classification, we will consider other classifiers and test it with a larger database in the future.

# References

[1] R. Khurana, M. Simons, J. F. Martin, and I. C. Zachary, "Role of angiogenesis in cardiovascular disease: a critical appraisal," *Circulation*, vol. 112, no. 12, pp. 1813–1824, Sep. 2005.

[2] C. Hetheridge, G. Mavria, and H. Mellor, "Uses of the in vitro endothelial-fibroblast organotypic co-culture assay in angiogenesis research," *Biochem. Soc. Trans.*, vol. 39, no. 6, pp. 1597–1600, Dec. 2011.

[3] C. Kanthou, Z. Gharaei, J. Haagen, S. J. Lunt, C. C. Reyes-Aldasoro, W. Doerr, and G. M. Tozer, "Abstract 5723: Inhibition of angiogenesis in the mouse heart by ionizing radiation," *Cancer Res.*, vol. 72, no. 8 Supplement, pp. 5723–5723, Apr. 2012.

[4] R. C. Gonzalez, S. L. Eddins, and R. E. (Richard E. Woods, *Digital image processing using MATLAB*, 2nd ed. [U.S.]: Gatesmark, 2009.

[5] "regionprops: Measure properties of image regions," *MATLAB R2013a online documentation*. [Online]. Available: http://www.mathworks.co.uk/help/images/ref/regionprops.html. [Accessed: 18-Aug-2013].

[6] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[7] "Ensemble Methods - Bag," *MATLAB R2013a online documentation*. [Online]. Available: http://www.mathworks.co.uk/help/stats/ensemble-methods.html#bsw8at7. [Accessed: 18-Aug-2013].

[8] C. C. Reyes-Aldasoro, "Retrospective shading correction algorithm based on signal envelope estimation," *Electron. Lett.*, vol. 45, no. 9, p. 454, 2009.

# Semi-automated cell counting in phase contrast images of epithelial monolayers

Rachel Flight[1] rxf144@bham.ac.uk
Gabriel Landini[2] g.landini@bham.ac.uk
Iain Styles[3] i.b.styles@cs.bham.ac.uk
Richard Shelton[2] r.m.shelton@bham.ac.uk
Michael Milward[2] m.r.milward@bham.ac.uk
Paul Cooper[2] p.r.cooper@bham.ac.uk

[1] PSIBS DTC, University of Birmingham, UK
[2] School of Dentistry, University of Birmingham, UK
[3] School of Computer Science, University of Birmingham, UK

**Abstract**

Non-destructive, quantitative evaluation of cell numbers in epithelial monolayers is necessary to monitor their proliferative rate in a single culture over a time period. Imaging using phase contrast microscopy is non-invasive and creates contrast in optically transparent cells; however intrinsic image artefacts mean that thresholding-based segmentation methods alone are insufficient to identify cells. The method described uses mathematical morphology to assign contrast to cells in a way that enables approximate segmentation by applying a threshold value. K-means clustering is subsequently applied to improve enumeration accuracy by removing segmentation artefacts. Cell numbers in images over multiple time-points may then be used to non-destructively generate growth curves. To validate this method we demonstrate a dose dependent relationship between H400 keratinocyte proliferation and concentration of foetal calf serum.

# 1 Introduction

Epithelial keratinocytes form stratified structures in regions of the body where their robust response to stress is essential, for example, in the masticatory mucosa. When the continuity of epithelial tissue is broken, keratinocytes respond by increasing proliferation and migrate into the wound to close it and prevent dehydration and infection [1]. Keratinocytes cultured *in vitro* form monolayers, which can be thought of as a simplified, 2D model of epithelium and are frequently used to study cell behaviour. The growth of such cultures in response to a stimulus may be assessed using cell counts obtained at multiple time-points.

The most widely-used method for determining cell number in cultures is by dissociating cells from their substrate, re-suspending and manually counting them on a haemocytometer using a microscope. The destructive nature of this technique means it is necessary to simultaneously maintain several cultures to generate growth curves. Small variations in the initial seeding densities of each culture or dilution upon re-suspension introduce error into the cell counts, therefore these are often performed at least in triplicate, which consumes both time and laboratory resources. Furthermore, operator variability contributes strongly to the count error [2].

Phase contrast (PC) microscopy creates contrast non-invasively in optically transparent objects using small changes in the phase of transmitted light so PC images of cell cultures

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

provide a potential approach for non-destructive cell counting. However, PC images are subject to intrinsic artefacts in the form of a 'halo effect': regions of high intensity pixels found at the edges of objects (Figure 1a). As a consequence, simple thresholding techniques are unsuccessful at distinguishing cells from background in PC images.

PC image analysis techniques frequently fall into two groups: textural analysis of entire images to distinguish between cell-populated and background regions, and high precision segmentation of individual cells. Methods in the first group are unable to provide information relating to individual cells and are not applicable at very low cell densities [3]. Conversely, methods in the second group are unsuitable for segmenting cells in dense images since they require time-consuming initialisation and high computational power [4]. Neither of these approaches is appropriate for both low and high density images of cell cultures, as is required to generate a comprehensive growth curve.

An alternative approach for locating cells in PC images involves correcting the halo artefact by using a mathematical model of PC optics to deconvolve images [5]. However this method is computationally expensive and therefore unsuitable for use on large images. A quicker, approximated form of deconvolution has been used to locate epithelial cells in scratch wound assays but to our knowledge has not been used to accurately count cells [6].

Lack of contrast in the related technique of brightfield microscopy makes it generally unsuitable for imaging unstained cells, but when the imaging plane is moved slightly above the focal plane, contrast increases such that cells become brighter than the background, albeit at the expense of subcellular detail. The opposite contrast is observed when the imaging plane moves below the focal plane, i.e. cells become darker. Dehlinger *et al.* located cells in a monolayer by subtraction of two such images acquired with the objective lens displaced by 15 µm [7]. However, this relatively small distance is difficult to achieve consistently without costly automated imaging apparatus.

In this paper we describe a method for emulating the contrast changes of defocusing brightfield microscopy by using mathematical morphology operators on a single, in-focus PC image. This produces an image in which cells are brighter than the background and thus may be segmented using a single threshold. We then perform k-means clustering on the segmented regions to identify and remove regions that do not correspond with cells and provide an accurate value for the number of cells in an image from which the total number of cells is extrapolated. This technique was applied to generate growth curves for H400 oral keratinocyte cultures supplemented with different concentrations of foetal calf serum.

# 2 Image analysis

## 2.1 Segmenting cells

When a mean filter is applied to a PC image using a circular kernel with radius, $r_k$, smaller than that of a cell, cells retain bright edges from the halo and a relatively dark centre (Figure 1b), analogous to the contrast observed by Dehlinger *et al.* when the imaging plane was focused below the focal plane. Conversely, when $r_k$ is larger than the average cell, cell centres become brighter than the margins (Figure 1c). Imposing a minimum pixel intensity of 0 and subtracting the small kernel image from the large kernel image generates an image in which cell containing regions are brightest (Figure 1d), and to which a single threshold may be applied using the Otsu method to locate cells (Figure 1e) [8].

For the purpose of selecting appropriate $r_k$values, a ground truth dataset was obtained by manually counting the cells in 10 images representing a range of cell densities. The

segmentation method described was applied using a range of radii for both large and small kernels to find the combination that minimised differences between this and the manual counts for all ground truth images.

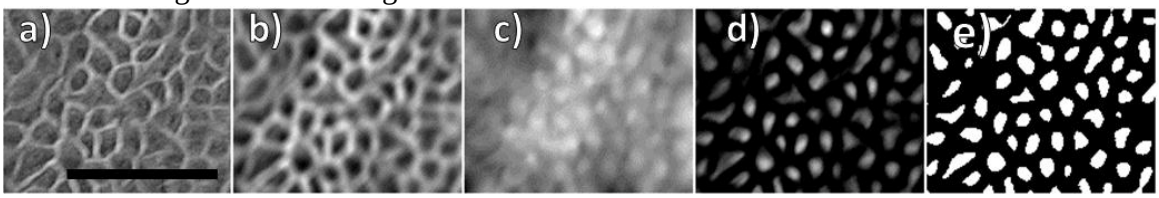



Figure 1: a) - Raw phase contrast image of H400 cells. b) - Image a) filtered with a small kernel mean filter, retaining bright cell margins and darker centres. c) - Original image a) filtered with a large kernel mean filter. Cell-containing regions are brighter than the surrounding regions. d) - Image b) subtracted from image c). Cell containing regions now appear brightest and Otsu thresholding is applied to obtain e). The contrast has been enhanced in all images for improved reproduction. Scale bar in a) is 150μm.

### 2.2 Classifying segmented regions

The cell count of an image is given by the total number of segmented objects in it. Image noise and edge effects at low cell densities can cause incorrectly segmented regions, which contribute erroneously to the cell count and must therefore be discounted. To this end, a dataset of 32 geometrical features describing the morphology and the greyscale properties of the corresponding pixels in the original image were calculated for each region using the "Particles8" plugin for ImageJ [9]. This feature dataset was subsequently reduced to 4 features by performing principal component analysis (PCA) on all segmented regions in all images acquired from a culture over the course of the experiment. K-means clustering (k=2) was applied using these 4 features to classify each segmented region as either "true" (real cell) or "false" (incorrect segmentation), and only "true" cells were counted.

### 2.3 Extrapolation of total cell count

The average count from a number of images was used to extrapolate the total number of cells in order to account for density differences across the culture; we used 20 images acquired at random locations in the culture. The fraction of the total culture area represented by one image was used to extrapolate total number of cells in the culture.

## 3 Experimental materials and methods

### 3.1 Cell cultures

Immortalised H400 keratinocytes derived from a human oral squamous carcinoma [10] were maintained at 37°C in a humidified atmosphere with 5% $CO_2$ in Dulbecco's MEM / nutrient mix supplemented with 10% foetal calf serum (FCS) (Biosera, UK), 0.6 μg/mL L-glutamine (Sigma, UK) and 0.4 mg/mL hydrocortisone (Sigma, UK).

For experimental analysis $3.6x10^4$ cells, as determined using a haemocytometer, were initially seeded into 35 mm culture dishes (Sarstedt, UK). To reduce the likelihood of non-uniform cell adhesion across the vessel due to uneven temperature distributions, culture dishes were preheated to 37°C prior to seeding and gently agitated following seeding. Duplicate cultures were supplemented with 2%, 6% and 10% FCS for analysis.

To enable comparison between cell counts obtained using the two approaches, additional cultures established under the same conditions were imaged and subsequently counted using an Improved Neubauer haemocytometer (as described below).

### 3.2 Image acquisition and analysis

A Nikon TE300 PC microscope with a x10 objective and a Nikon D40 digital camera (Nikon, UK) were used to acquire images at 2000 x 3008 pixels. Conversion from pixels to mm was obtained via calibration with an image of a microscope stage micrometer giving an image size of 1.17 x 0.78 mm. Images were acquired between 48 - 120 hours post-seeding. The optimal kernel radii for this experimental set-up were determined to be 80 pixels (31 μm) and 14 pixels (5 μm). Cell segmentation and classification steps were implemented in ImageJ and included the use of freely available "Morphological_PCA" and "Morphological_Clustering" macros [9][11]. Computation time on a set of 280 images was 10 minutes using an Intel® Core™ i7-3770 processor. The kernel radius selection method was implemented using Matlab version R2013a [12].

### 3.3 Cell enumeration using haemocytometer method

Immediately following imaging, the media was aspirated from the culture vessel and replaced with 1 mL trypsin, and incubated until cells were detached. For ease and accuracy of counting at high cell numbers, the cell suspension was diluted appropriately with culture media. Cell concentration was measured using an Improved Neubauer haemocytometer by counting the number of cells in each of the nine 1 $mm^2$ squares (representing a volume of 0.1 μL each) for each of the two haemocytometer chambers and averaging to provide the number of cells per 0.1 μL. To extrapolate the total number of cells in the culture this value was multiplied by $10^4$ and the dilution factor dependent upon the amount of media added.

### 3.4 Assessing classification success

The success of k-means classification was assessed by comparison with 20 images across all samples in which all segmentations had been labelled manually. For this comparison, true positive (TP) describes segmented cells correctly identified by the clustering method described, and true negative (TN) describes erroneous segmentations correctly classified as such. False positive (FP) describes artefacts wrongly classified as cells and false negative (FN) describes cells classified as artefacts. The $F_1$-score is a combined measure of precision and recall, for which 0 indicates no agreement between manual and k-means classifications and 1 indicates perfect agreement. It is calculated as:

$$F_1 = 2\left(\frac{pr}{p+r}\right) \quad (1)$$

where $p = \frac{TP}{TP+FP}$ is precision and $r = \frac{TP}{TP+FN}$ is recall. The number of cells undetected in the segmentation step was also recorded for these 20 images.

# 4 Results

The average $F_1$-score of segmentation classification using k-means clustering was 0.94 ± 0.04 versus 0.86 ± 0.06 without classification. Average precision was 0.92 ± 0.06 and average recall was 0.96 ± 0.03.

Figure 2a shows growth curves obtained using the method described, which indicate that H400 keratinocytes proliferate in a dose dependent manner with FCS concentration.

Figure 2b shows the growth curves obtained from a set of cultures using the image analysis method described here and the more routinely used haemocytometer method. The image analysis method was systematically lower than the haemocytometer-measured count, although within one standard deviation. The high $F_1$-score indicated that this discrepancy was not likely due to errors in cell classification. On average, 3.8 ± 1.4% of cells per image were undetected in the segmentation step, but this does not fully account for the difference in counts between techniques, which is 22 ± 5%. The systematic operator error associated with haemocytometer counting could potentially have contributed to the discrepancy. Further work is required to determine the source of this difference.

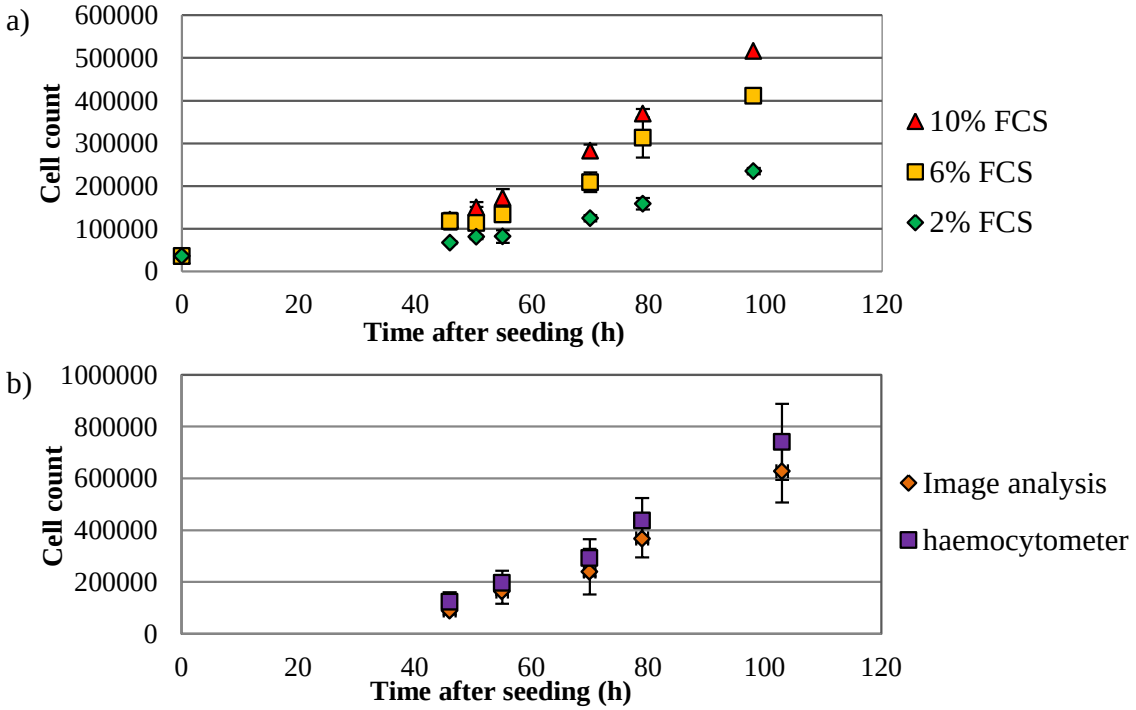



Figure 2: a) Growth curves for H400 keratinocyte cultures supplemented with different concentrations (2-10%) of FCS calculated using the method described. Growth rates were dose-dependently associated with FCS supplementation. Error bars indicate standard deviation between the average cell counts of two cultures under the same culture conditions. b) H400 keratinocyte growth curve determined using two enumeration methods. Each time point corresponds with a single culture from which the total number of cells was determined initially using the image analysis method described and subsequently using the haemocytometer approach. Error bars indicate ±1 standard deviation of the mean.

# 5 Discussion and conclusion

The most significant advantage of the method described compared with other commonly used cell counting approaches is that it is non-destructive and thus cultures may be counted multiple times. Aside from minimising human errors associated with the haemocytometer method and reducing the time and cost of maintaining many cultures to generate a growth curve, this image analysis method also introduces the opportunity for the study of the spatial distribution of cells under different culture conditions.

Image filtering has been used previously to locate cells in monolayers, for example, Sarsby *et al.* used a difference of Gaussians filter in their scratch wound analysis

algorithm[6]. A key difference in our approach is the parameterisation method, which prioritises a single region being segmented per cell, to reduce the likelihood of cells merging in the segmented image; essential given our aim of accurate cell counting. By using classification with the unsupervised k-means clustering technique, incorrect segmentations were accurately identified and removed without requiring the use of time-consuming manual training sets. However, it may be possible to further improve accuracy using supervised classification techniques and this will be investigated in further work.

We have shown that our method may be used to distinguish between the growth profiles of keratinocyte monolayers established in different culture conditions. The cell counts obtained by this method are within one standard deviation of that obtained from the same culture using a haemocytometer, although systematically lower. Further work is required to determine the source of this difference.

# Acknowledgements

The authors thank Ela Claridge for useful initial discussions and the EPSRC for funding through a studentship to RF in the PSIBS Doctoral Training Centre (EP/F50053X/1).

# References

[1] P. Martin, "Wound Healing--Aiming for Perfect Skin Regeneration," *Science*, 276(5309):75–81, 1997.

[2] R. Biggs and R. L. Macmillan, "The error of the red cell count," *Journal of clinical pathology*, 1(5):288–291, 1948.

[3] N. Jaccard, L. D. Griffin, A. Keser, R. J. Macown, A. Super, F. S. Veraitch, and N. Szita, "Automated method for the rapid and precise estimation of adherent cell culture characteristics from phase contrast microscopy images.," *Biotechnology and bioengineering*, 111(3):504–517, 2013

[4] I. Seroussi, D. Veikherman, N. Ofer, S. Yehudai-Resheff, and K. Keren, "Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes.," *Journal of microscopy*, 247(2):137–146, 2012.

[5] Z. Yin, T. Kanade, and M. Chen, "Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation.," *Medical Image Analysis*, 16(5):1047–1062, 2012.

[6] J. Sarsby, E. Claridge, G. Nash, D. Thickett, H. Jeffery, and S. Zheng, "Restoration of phase contrast microscopy images for the analysis of lung epithelial scratch wound repair assays," in *MIUA* 2012, 101–106.

[7] D. Dehlinger, L. Suer, M. Elsheikh, J. Peña, and P. Naraghi-Arani, "Dye free automated cell counting and analysis.," *Biotechnology and bioengineering*, 110(3):838–847, 2013.

[8] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Sytstems, Man and cybernetics*, 20(1):62–66, 1979.

[9] G. Landini, "Advanced shape analysis with ImageJ," in Proceedings of the Second ImageJ user and developer Conference, 2008, 116–112, http://www.mecourse.com/landinig/software/software.html

[10] S. Prime, S. Nixon, and I. Crane, "The behaviour of human oral squamous cell carcinoma in cell culture," *The Journal of Pathology*, 160(3): 259–269, 1990.

[11] W. S. Rasband, "ImageJ." U. S. National Institutes of Health, Maryland, USA., 1997-2014. http://imagej.nih.gov/ij/

[12] The MathWorks Inc., "Matlab." Natick, MA.

# A Wavelet-based Approach for Analyzing Non-Stationary Phonation Sequences from Endoscopic High-Speed Videos

Jakob Unger[1]
unger@hochschule-trier.de
Maria Schuster[2]
Dietmar Hecker[3]
Bernhard Schick[3]
Joerg Lohscheller[1]

[1] Dept. of Computer Science
Trier University of Applied Sciences
Trier, Germany

[2] Dept. of Audiology and Phoniatrics
University Hospital Munich
Munich, Germany

[3] Dept. of Otorhinolaryngology
Saarland University Hospital
Homburg/Saar, Germany

**Abstract**

Direct observation of vocal fold dynamics is an essential part of clinical diagnosis of voice disorders. Presently, high-speed videoendoscopy is a state-of-the-art procedure to capture the vibratory behaviour of the vocal folds in full images. In order to enable an objective assessment of high-speed video sequences, a wavelet-based method for extracting descriptive features from phonovibrograms, a two-dimensional image containing the spatio-temporal pattern of vocal fold dynamics, was presented. The method quantifies basic characteristics of vocal fold dynamics that are closely related to a basic set of video-based measurements categorized by the European Laryngological Society for a subjective assessment of pathologic voices. Here, features are quantified from sustained phonation at habitual pitch and loudness. Stationary voice production mechanisms, however, do not fully reflect the complex mechanisms of voice production as a whole. Therefore, pathologies might not be predicted reliably. This study addresses an extension of the wavelet-based procedure to non-stationary phonation sequences. The enhanced discriminative power is demonstrated by healthy and pathologic subjects that were examined during sustained and non-stationary phonation.

## 1 Introduction

The voice production mechanism is primarily given by the vocal folds (VF) modulating the passing flow of air. Pathologic voice is commonly caused by irregular or asymmetric vibratory behaviour [2]. Clinical diagnostics therefore demands visual inspection of the VFs during phonation. The rapid movement of the VFs is recorded with endoscopic high-speed video cameras capturing thousands of frames for each second. In order to facilitate a comprehensive documentation of laryngeal dynamics, phonovibrograms (PVG) [5] were introduced

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

that allow -besides visualizing even long high-speed videos in compact graphs- also a computerized and hence, objective analysis of VF vibration. From the PVG, a limited number of clinical relevant parameters can be derived by employing a wavelet-based analysis [8]. The wavelet-based analysis characterizes vibration patterns examined during sustained phonation with constant pitch and loudness. The feature's discriminative power has been evaluated for different pathologic findings [7] and showed a good performance so far. However, the VFs' vibration patterns strongly depend on pitch and intensity settings [9]. Therefore, the vibration during sustained phonation characterizes merely a part of the complex voice production mechanism. First studies using non-stationary phonation and a model-based approach showed promising results [9]. Defining adequate models, however, remains extremely challenging. In the current study, we extended the wavelet-based approach to non-stationary phonation sequences. It is shown that healthy and pathological vibration patterns can be differentiated more reliably when analysing non-stationary phonation sequences.

# 2 Materials and Methods

## 2.1 Subjects and Equipment

Laryngeal high-speed recordings were performed for a total number of 10 female subjects. For five subjects no signs of voice disorders were found whereas the other five subjects were with a diagnosed unilateral VF paresis. Each subject was examined twice, during sustained phonation at habitual pitch and loudness and during a pitch raise from low to high pitch. The recordings were made with the HS Endocam 5562 highspeed camera system (Richard Wolf GmbH, Knittlingen, Germany) providing a sampling rate of $4,000$ frames per second and a spatial resolution of $256 \times 256$ pixels.

## 2.2 Phonovibrography

In order to quantify the VFs' lateral movement along the visible anterior-posterior dimension the glottal area, enclosed by left and right VFs, is segmented in each video frame using a modified region growing algorithm (Fig. 1 (a) ) that proved to provide reliable segmentation results even for low image quality [4]. The segmentation is then transformed to the phonovibrogram (PVG) visualization [5] that encodes the time varying deflection of both VFs from the glottal midline in a single colour image (Fig. 1 (b) and (c) ). Besides being a valuable visualization for a subjective analysis of laryngeal function the PVG provides the basis for image processing routines to classify different types of pathologic vibration patterns.

## 2.3 Wavelet-based Analysis of Phonovibrograms

The PVG exhibits periodically occurring geometric structures for each VF, respectively. An example is illustrated in Figure 1 (c). Here, the triangular shape (white dashed line) represents a "zipper-like" opening and closing. In order to quantify the periodical structures, the wavelet phase is estimated using a complex Morlet wavelet $\psi_1$. Let $g_{l,r}$ denote the area enclosed by the left or right VF and the glottal axis (Fig. 1 a). The phase of left and right VF vibration reads as

$$\phi_{l,r}(b) = \arctan \frac{\Im\{\mathscr{W}_{\psi_1}\{g_{l,r}\}(a^{l,r}(b),b)\}}{\Re\{\mathscr{W}_{\psi_1}\{g_{l,r}\}(a^{l,r}(b),b)\}} \tag{1}$$

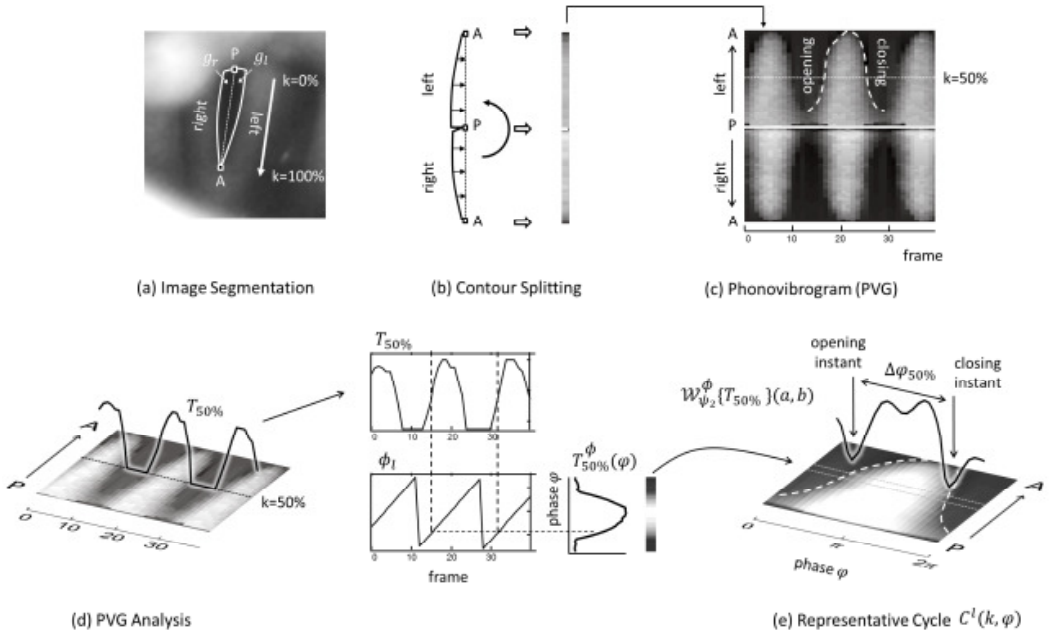



Figure 1: (a), (b), (c) Phonovibrogram construction process, (d), (e) Wavelet-based PVG analysis

where $\mathscr{W}$ denotes the wavelet transform and $a^{l,r}(b)$ denotes the wavelet scale chosen for a specific frame $b$. In [8], a constant scale $a(b) = a_0$ was taken assuming a constant pitch during laryngoscopy. In order to evaluate non-stationary phonation, in this study, the scale $a(b)$ is chosen from the wavelet ridges as proposed by Carmona et al. [1]. A mean PVG cycle can then be obtained by assigning the PVG values $T_k^{l,r}$ at position k along the anterior-posterior axis to the corresponding phase (Fig. 1 (e) ).

$$C^{l,r}(k,\varphi) = \sum_{M=\{b|\phi_k^{l,r}(b)=\varphi\}} \frac{1}{|M|} T_k^{l,r}(b) \tag{2}$$

The graph $C^{l,r}$ shows the dominant geometric structure periodically appearing within the PVG representation. This pattern is essentially determined by the opening and closing instants forming the triangular shape. To localize these instants a second order Gaussian wavelet $\psi_2$ is employed. The wavelet has two vanishing moments and hence, acts as differential operator of the smoothed source signal [6] providing minima at opening and closing instants (Fig. 1 f). A more precise localization of these instants can be achieved by employing a multiscale product (MSP) as done by Unger et al. [8] involving the product of several successive wavelet scales. Thereby, opening and closing instants are indicated by sharp peaks whereas noise components are suppressed.
Finally, in order to separate opening and closing instants from each other a first order Gaussian wavelet $\psi_3$ is employed. Consequently, the sign of the wavelet transform

$$\operatorname{sgn}\left(\mathscr{W}_{\psi_3}\{C^{l,r}\}(a,b)\right) \tag{3}$$

indicates whether the MSP minimum refers to glottal opening or closing.

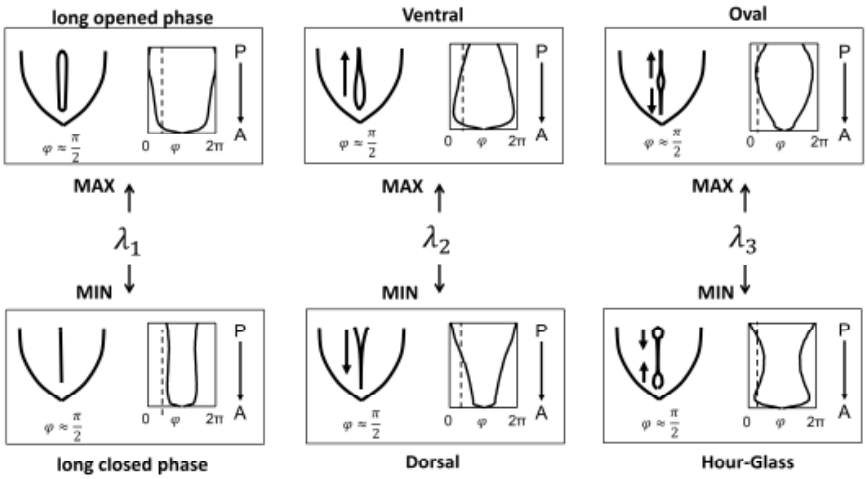



Figure 2: Correlation of the eigenvalue decomposition and VF vibration patterns.

## 2.4 Principal Component Analysis

A compact description of the contours is achieved by computing the distance from opening to closing instants $\Delta\varphi_k$ along the entire length of the glottal axis (Fig. 1 (f) ). From the distance vector $\Delta\varphi_k$, a PCA is performed for 100 healthy subjects during sustained phonation. The PCA space spanned by these subjects serves as norm-map where the remaining subjects to be analysed are projected into. The eigenvectors spanning the PCA space and examples of corresponding glottal configurations are depicted in Fig. 2. The first eigenvector exhibits a rectangular shape affecting the contour's width in phase direction. A long and complete closure is encoded by low $\lambda_1$ values whereas short or incomplete closure corresponds to high $\lambda_1$ values. Triangular geometries comply with a zipper-like opening and closing. Opening from anterior to posterior is given by high $\lambda_2$ (ventral) and vice versa, opening from posterior to anterior corresponds to low values of $\lambda_2$ (dorsal). Values of $\lambda_2$ around zero can be interpreted as superposition of both types meaning that the VFs open and close simultaneously alongside the glottal axis. Finally, the third eigenvector encodes oval and hour-glass contour shapes meaning opening from medial to terminal and vice versa, from terminal to medial. The first three eigenvectors make up more than 90% of the overall variance and hence, comprehensively describe vibration patterns for each VF individually. Furthermore, the first three eigenvalues are closely related to a categorical scheme for a subjective assessment of pathologic voices defined by the European Laryngological Society [3] demonstrating the clinical interpretability of the eigenvalue decomposition.

## 2.5 Assessment of non-stationary Phonation Sequences

By projecting a PVG into the given PCA space the vibration characteristics can be assessed with merely three parameters for each VF. Furthermore, when taking windowed PVG segments, non-stationary phonation sequences will form a trajectory through the PCA space comprehensively describing alterations of vibration characteristics over time. In the current study, windowed segments with a length of 400 frames ($= 0.1sec$) were shifted by 200 frames in each time step for the non-stationary recordings whereas the entire length of the signal was taken for the stationary sequences.

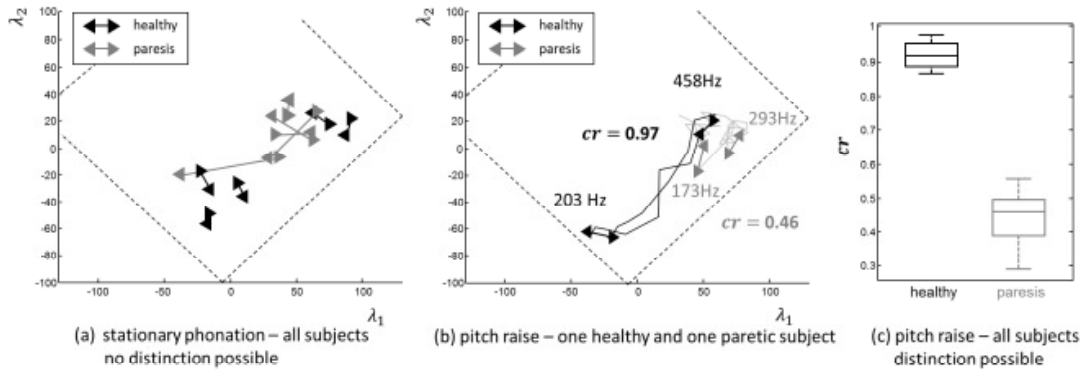



Figure 3: (a), (b) PCA space characterizing VF vibration patterns during pitch raise from low to high frequency, (c) Linear correlation of left and right vibration patterns during pitch raise.

In order to evaluate whether the vibration patterns of left and right VF are changing synchronously, the correlation coefficients $c_{\lambda_i}$ between the eigenvalues $\lambda_i$ of left and right VF vibration patterns are computed. Let $s_{\lambda_i}$ denote the range of the $i^{\text{th}}$ eigenvalue paramter (during pitch-raise) and $s_0 = \sum_{i=1}^{3} s_{\lambda_i}$ . In this respect, the measure

$$cr = \frac{s_{\lambda_1}}{s_0}\left|c_{\lambda_1}\right| + \frac{s_{\lambda_2}}{s_0}\left|c_{\lambda_2}\right| + \frac{s_{\lambda_3}}{s_0}\left|c_{\lambda_3}\right| \tag{4}$$

quantifies the weighted linear correlation between left and right vibration patterns during pitch raise.

# 3 Results and Discussion

Figure 3 (a) shows the projections of left and right VF vibration patterns of all subjects examined during sustained phonation. A higher asymmetry can be seen for the paretic subject illustrated by the higher distance between the projections of left and right VF within the PCA space. Although asymmetry is obviously higher for the paretic group, no clear distinction between healthy and paretic subjects can be made from the eigenvalue measures alone.
Figure 3 (b) exemplarily demonstrates the change of the vibration pattern during a pitch raise for a healthy and a paretic subject. For the healthy subject, a systematic change from dorsal zipper-like vibration to a simultaneous opening and closing (longitudinal closure) accompanied by a gradually increasing opened-closed ratio can be seen. Contrarily, no systematic change can be observed for the paretic subject. For the entire sequence, high $\lambda_1$ values are observed for the paretic subject corroborating the fact that paresis is often accompanied by a complete absence of VF contact. For the healthy subject high $\lambda_1$ values are merely seen for higher pitches. Furthermore, the correlation measure $cr = 0.46$ indicates a reduced coupling between both VFs compared to the healthy subject ($cr = 0.97$). A boxplot of the correlation measure $cr$ evaluated for all subjects is given in Figure 3 (c). From this measure, healthy and paretic subjects can be separated adequately.

# 4 Conclusion

Analysis of non-stationary phonation provides additional information of the underlying laryngeal dynamics that cannot be derived from stationary processes. It has been shown that additional clinical relevant measures can be obtained providing valuable information for a diagnosis of voice disorders. In order to verify these advantages more quantitative measures have to be defined and their relevance for different pathologic findings have to be clarified in more extensive studies.

# 5 Acknowledgement

This work is supported by the German Research Foundation (DFG), Grant No. Lo-1413/2-2.

# References

[1] R.A. Carmona, W.L. Hwang, and B. Torresani. Multiridge detection and time-frequency reconstruction. *IEEE Trans Signal Processing*, 47(2):480–492, 1999.

[2] U. Eyshold, F. Rosanowski, and U. Hoppe. Irregular vocal fold vibrations caused by different types of laryngeal symmetry. *Eur Arch Otorhinolaryngol*, 260(8):412–417, 2003.

[3] G. Friedrich and P. H. DeJonckere. The voice evaluation protocol of the european laryngological society (els) – first results of a multicenter study. *Laryngo Rhino Otol*, 84(10): 744–752, 2005.

[4] J. Lohscheller, H. Toy, U. Eysholdt, and M. Doellinger. Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos. *Med Image Anal*, 11(4):400–413, 2007.

[5] J. Lohscheller, U. Eysholdt, H. Toy, and M. Doellinger. Phonovibrography: Mapping high-speed movies of vocal fold vibrations into 2-d diagrams for visualizing and analyzing the underlying laryngeal dynamics. *IEEE Trans Med Imag*, 27(3):300–309, 2008.

[6] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, ISBN: 0123743702, 2009.

[7] J. Unger, M. Schuster, D.J. Hecker, B. Schick, and J. Lohscheller. A multiscale product approach for an automatic classification of voice disorders from endoscopic high-speed videos. In *Engineering in Medicine and Biology Society (EMBC)*, pages 7360–7363, 2013.

[8] J. Unger, D.J. Hecker, M. Kunduk, M. Schuster, B. Schick, and J. Lohscheller. Quantifying spatiotemporal properties of vocal fold dynamics based on a multiscale analysis of phonovibrograms. *IEEE Trans Biomed Eng, in press*, 2014.

[9] T. Wurzbacher, R. Schwarz, M. Doellinger, U. Hoppe, U. Eysholdt, and J. Lohscheller. Model-based classification of nonstationary vocal fold vibrations. *J Acoust Soc Am*, 120 (2):1012–1027, 2006.

# 3-D MRI Brain Scan Classification of Epilepsy Versus Non-epilepsy

Akadej Udomchaiporn[1]
akadej@liv.ac.uk
Frans Coenen[1]
coecen@liv.ac.uk
Vanessa Sluming[2]
vanessa.sluming@liv.ac.uk
Marta García-Fiñana[3]
m.garciafinana@liv.ac.uk

[1] Department of Computer Science
University of Liverpool
[2] School of Health Science
University of Liverpool
[3] Department of Biostatistics
University of Liverpool

**Abstract**

A 3-D classification method is presented for application to Magnetic Resonance Imaging (MRI) brain scan volumes. The objective of the classification is to identify volumes that feature indicators of epilepsy against volumes that do not, by considering the shape and size of the lateral (left and right) ventricles of the brain. The dataset used with respect to the reported experimentation comprised 210 3-D MRI brain scans of which 105 were from epilepsy patients and the remainder from healthy people. The classification is supported by two proposed point series based representation techniques: (i) Disc and (ii) Spoke. The reported experimental results indicated that our proposed techniques can classify 3-D MRI brain scan data with a classification accuracy of up to 69.81%.

## 1 Introduction

Automated (or semi-automated) medical prediction is a challenging real world problem. The effective and efficient automated prediction of medical conditions is clearly of significant benefit especially with respect to the use of resources, even if used as a "first screening". The work presented in this paper is directed at the automated screening for epilepsy from Magnetic Resonance Imaging (MRI) brain scan volumes (although the work has clear application elsewhere). Epilepsy is a medical condition whereby nerve cell activity in the brain is disturbed; it causes abnormal behaviour accompanied by symptoms such as loss of consciousness or convulsions. There are various indicators for the cause of epilepsy, one is the shape and size of the ventricles contained in the brain [7]. This can be measured by manual inspection of MRI scan data. MRI brain scan data consists of a sequence of 2-D "slices" in three planes: Sagittal (SAG), Coronal (COR), and Transverse (TRA). Collectively we refer to this set of slices as a *volume*. However, although tools exist to support manual inspection (for example the Brain Voyager range of software products), the process is time consuming, challenging and error prone. Automation of the process therefore seems desirable.

©2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

This paper proposes the use of a machine learning approach, more specifically a classification approach. Classification is concerned with the automated allocation of labels to data using a piece of software known as a classifier; in our case the input is a description of the lateral ventricles and the labels are the set $\{epilepsy, \neg epilepsy\}$. However, the classification process also entails a number of challenges, the most significant of which are: (i) the production of the classifier and (ii) the mechanism for representing the input in such a way that the key information is retained while at the same time ensuring tractability (the two challenges are related). Classifiers are typically generated using pre-labelled training data. There are various techniques whereby classifiers can be learnt, in this paper we explore the use of two. So as to obtain a level of confidence in a generated classifier it is typically evaluated against pre-labelled test data.

The representation of the input data is a key part in the classifier generation process. The image representation will affect both the efficiency and effectiveness of the machine learning. In this paper, two representation techniques, referred to as the Disc-based and Spoke-based representations, are proposed. Both are point series representation techniques that are designed to capture the boundary a Volume Of Interest (VOI) in terms of a series of points which in turn can be conceptualised as a curve in 2-D space. Such curves can be used directly for classification purposes using (for example) a *k*NN (*k* Nearest Neighbour) process [3] which requires some form of similarity measure; in this paper we use the length of the shortest "warping path" generated using the well established Dynamic Time Warping (DTW) technique [1]. Alternatively the curves can be processed further so that a *feature space* model is derived which will allow for the individual VOI to be represented in terms of a *feature vector*. In this paper we suggest the use of Hough "signature" extraction for this purpose. Feature vector representations are compatible with a number of "standard" classifier generation models. In this paper the Support Vector Machine (SVM) [2] model is used but any other form of classifier generation model may be equally well applicable. Both the Disc-based and Spoke-based representations are described as well as the *k*NN and SVM classification processes, together with a full evaluation.

## 2 Point Series Generation

From the above, two techniques for generating a point series representation are proposed, the Disc-Based technique and the Spoke-Based technique. The input in both cases is a previously segmented description of the VOI (the lateral ventricles in our case). With respect to the evaluation presented later in this paper the authors used the *Bounding-Box* segmentation technique proposed and described in [10] together with a thresholding technique and noise removal. Other similar techniques would have sufficed however experiments presented in [10], indicate that the Bounding-Box technique works well with respect to the lateral ventricles. Essentially the segmentation results in a *zero-one* representation where voxels that are part of a VOI (the lateral ventricles) are marked with a 1 (black) and voxels that are not part of the VOI are marked with a 0 (white). The Disc-based technique is described in Sub-section 2.1 and the Spoke based technique in Sub-section 2.2 below.

### 2.1 Disc-based Representation Technique

The Disc-based representation technique is illustrated in Figure 1. A point series is generated by taking a series of measurements from the centroid of the VOI (the lateral ventricles)

to each voxel located on the boundary of the VOI. This is achieved by iteratively moving a 2-D plane along the length of the VOI voxel by voxel along a selected axis. At each iteration the intersection between the plane and the VOI was identified and measurements taking from the centroid of the VOI to the the voxels located along the boundary of the intersection. With respect to the ventricles the intersection typically took the form of a "disc" shape, hence the name "Disc-based" representation. Whatever the case the measurements were then used to describe a 2-D curve with distance along the Y-axis and the sequential measurement identification number along the X-axis. In the case of the ventricles three such curves were generated one for each of the the three cardinal axes: Sagittal, Coronal and Transverse. A perceived disadvantage of the Disc-based technique is that a large number of points are generated giving rise to substantial curves which in turn might require considerable computational resource to process (it is desirable for screening to be conducted at time of consultation, thus in real time using the processing power available on standard desktop machines). Note also that the Disc-based technique, given a collection of segmented ventricles, will produce curves of different lengths (ventricles will be of different sizes) which is not ideal for comparison purposes. The Spoke-based technique described in the following sub-section seeks to address this issue.

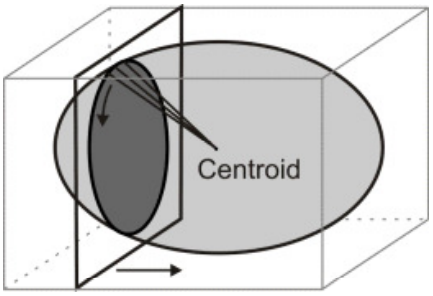



Figure 1: Disc-based point series generation

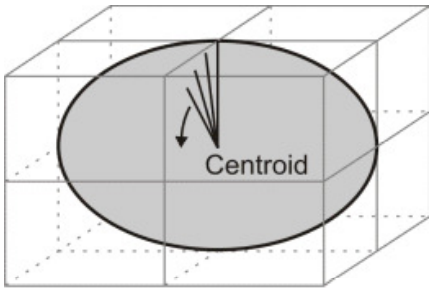



Figure 2: Spoke-based point series generation

## 2.2 Spoke-based Representation Technique

The Spoke-based point series generation technique is illustrated in Figure 2. The idea is to measure the distance from the centroid to the VOI boundary only in the three cardinal planes (Sagittal, Coronal and Transverse in the case of our ventricles). This is achieved by considering each plane in turn and conceptually generating a set of "spokes" radiating out from the centroid, spaced at some angle of separation $x^{\circ}$, and then measuring the distance from the centroid to where the spoke cuts the boundary of the volume. The results again three point series, but in this case each point series contains fewer points than in the case of the Disc-based technique. In addition the curves will all be of the same length provided the same value of $x$ is used through out.

# 3 Classification

For classification purposes we require a training set of pre-labelled data (in our case a MRI brain scan dataset with each scan labelled as either $epilepsy$ or $\neg epilepsy$). Point series (curves) can then be generated from the data and a label associated with each curve. These curves can then be used either directly or for classification purposes. In the first case classification can be conducted simply be comparing new curves associated with unlabelled VOI

with existing curves in our "curve base". The proposed process is described in sub-section 3.1. In the second case we translate the point series into a standard feature space representation compatible with many "off-the-shelf" classifier generators. This process is described in sub-section 3.2.

### 3.1 Direct Classification

Using the generated point series (curves) directly to classify "unseen" data the use of $k$-Nearest Neighbours ($k$NN) classification [3] is proposed with $k$ set to 1. The $k$NN classification approach requires a distance measure, Euclidean distance is frequently used for this purpose, however this is clearly unsuitable when comparing curves (which may also be of different length). Instead Dynamic Time Warping (DTW) as described in [1] was used to identify a *warping path*. DTW is a well established technique used to compare curves. DTW operates as follows. Given two curves, $X$ with length $m$ and $Y$ with length $n$, a matrix $A$ is constructed with $m$ rows and $n$ columns. Each element $(i, j)$ within matrix $A$ describes the distance between point $i$ on curve $X$ and the point $j$ on curve $Y$. The goal is to find the "warping path" through this matrix describing the shortest distance from $(0,0)$ to $(m,n)$. To improve the efficiency with which this warping path can be identified the "Sakoe-Chiba" band [9] was used to define a "constraint region". The length of the warping path can then be used as a measure of the similarity between two curves. Using this approach the most similar curve in the curve base to a given new curve (representing a previously unseen ventricle extracted from a MRI brain scan) can be identified and the label from the identified curve used to label the new curve.

### 3.2 Classification After Further Processing

The idea behind the further processing of the point series data is that this might result in a more effective representation for classification purposes in that redundant information may be removed. More specifically the idea is to generate a collection of point series using a pre-labelled training set and then converting these point series into a set of "feature vectors" or "signatures" representing key elements of the point series. Collectively the feature vectors (signatures) describe a multi-dimensional feature space (one dimension per feature where each feature has a number of potential values associated with it) in which each point series is described by a single location within the space described by the vector from the origin of the feature space to the location. To this end Hough feature extraction, based on the Hough concept [6], was used to identify signatures from within the point series (curve) data. The curves were firstly transformed into a parameter space (accumulator matrix) $A$ comprised of $m$ rows and $n$ columns where $m$ is the difference between the minimum and maximum collected distance and $n$ is the total number of points in the point series. Note that the value of $m$ varied with respect to different point series. The process, based on the Hough transform, described in [11] was then used to extract signatures (feature vectors) from the accumulator matrix. Feature vector representations are compatible with many standard forms of classifier generator. With respect to the evaluation described in the following section a SVM classification model [2] was used.

## 4 Evaluation

From the foregoing two point series representation methods are proposed (Disc-based and Spoke-based) which can be either used directly for classification purposes (using $k$NN cou-

pled with DTW in our case) or processed further and then used for classification purposes (using Hough signature extraction and the SVM classification model in our case). The comparative evaluation of these four different mechanisms is presented in this section. For the evaluation a dataset comprised of 210 MRI brain scans, obtained from the Magnetic Resonance and Image Analysis Research Centre at the University of Liverpool, was used. Each scan consisted of 256 two dimensional (2-D) parallel image slices in each of the three cardinal planes. The resolution of each image slice was 256 x 256 pixels (voxels) with colour defined using 8-bit gray scale (256 colours). The dataset is described in more detail in [10]. It should also be noted that it has been used in a number of other studies, for example in [4].

The reported experiments were conducted using Ten-fold Crossed Validation (TCV) where different tenths of the data were used as the test set. In the case of the Spoke-based technique a sequence of values for $x$ was used ($\{1°, 2°, 3°, 4°\}$). The SVM model provided with The Waikato Environment Knowledge Analysis (WEKA) data mining workbench [5] was used where required. The classification results obtained are shown in Tables 1 and 2. Table 1 shows the results obtained using direct classification ($k$NN and DTW similarity measurement) and Table 2 the results after further processing (Hough signature extraction and SVM). Both tables list the accuracy (Accu.), sensitivity (Sens.) and specificity (Spec.) values obtained. From the tables it can be observed that: (i) the Spoke-based representation outperformed the Disc-based representation, (ii) the Spoke-based representation tended to improve as the value of $x$ increased and then drop off (using direct classification $x = 2°$ tended to produced the best results, while with further processing $x = 3°$ tended to produced the best results) and (iii) direct classification produced better overall average results than obtained after further processing. The reason for the Spoke-based representation outperforming the Disc-based representation is conjectured to be because the Spoke-based representation was more succinct and therefore less cluttered (less room for ambiguity). The conjectured reason for the direct classification approach outperforming the alternative approach is that the additional processing conducted in the later case had the effect of coarsening the representation with the result that some information was lost. The best accuracy, sensitivity and specificity values were all obtained using the Spoke-based representation and direction calcification: 69.81% ($x = 2°$), 75.47% ($x = 3°$) and 67.92% ($x = 2°$) respectively. It should also be noted that the Spoke-based techniques was also more efficient because the resulting point series encompassed fewer points.

| Technique | Accu. | Sens. | Spec. |
|---|---|---|---|
| Disc | 62.20 | 67.50 | 57.14 |
| Spoke ($x = 1°$) | 64.15 | 66.04 | 62.26 |
| Spoke ($x = 2°$) | **69.81** | 71.70 | **67.92** |
| Spoke ($x = 3°$) | 68.87 | **75.47** | 62.26 |
| Spoke ($x = 4°$) | 60.98 | 67.50 | 57.14 |
| Average | 65.20 | 69.64 | 61.34 |

Table 1: Classification results obtained using $k$NN and DTW

| Technique | Accu. | Sens. | Spec. |
|---|---|---|---|
| Disc | 59.43 | 58.49 | **60.38** |
| Spoke ($x = 1°$) | 62.20 | 67.50 | 57.14 |
| Spoke ($x = 2°$) | **64.63** | **70.00** | 59.52 |
| Spoke ($x = 3°$) | 61.32 | 62.26 | **60.38** |
| Spoke ($x = 4°$) | 58.49 | 62.26 | 57.42 |
| Average | 61.21 | 64.10 | 58.97 |

Table 2: Classification results obtained using SVM and Hough Signature Extraction

## 5 Conclusions

This paper has proposed an approach to 3-D MRI brain scan classification using two point series based representations, Disc-based and Spoke-based. These can be used directly (a $k$NN based approach is suggested) or processed further (a Hough signature extraction approach

is suggested) and then input to a standard classifier generation model (the SVM model was used with respect to the reported evaluation). The approaches were evaluated in the context of epilepsy classification with respect to lateral ventricle data obtained from 3-D MRI brain scans. The main findings were that the Spoke-based representation outperformed the Disc-based representation (especially when $x = 2^\circ$ or $x = 3^\circ$) and that direct classification produced better results than when further processing was applied to the point series data. Compared to other previous works the results reported in [4] were slightly better than those reported here, although this work was directed at the corpus callosum which might be a better indicator of epilepsy. Better results were also reported in [8], using an oct-tree representation of the ventricles, although in this case the work was directed at classifying Alzheimer's disease and level of education rather than epilepsy.

# References

[1] R. Bellman and R. Kalaba. On Adaptive Control Processes. *IRE Transactions on Automatic Control*, 4(2), 1959.

[2] C. Cortes and V. Vapnik. Support-vector Networks. *Machine Learning*, 20(3):273–297, 1995.

[3] G. Dougherty. *Digital Image Processing for Medical Applications*. Cambridge Press, 2009.

[4] A. Elsayed, M. Hijazi, F. Coenen, M. García-Fiñana, V. Sluming, and Y Zheng. Time series case based reasoning for image categorisation. In *Case-Based Reasoning Research and Development*, pages 423–436. Springer, 2011.

[5] M. Hall, E. Frank, and G. Holmes. The WEKA Data Mining Software: An Update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.

[6] P. V. C. Hough. Method and Means for Recognizing Complex Patterns, 1962. URL http://www.google.com/patents/US3069654. US Patent 3,069,654.

[7] J.L. Lin, M. Mula, and B.P. Hermann. Uncovering the Neurobehavioural Comorbidities of Epilepsy over the Lifespan. *The Lancet*, 380(9848):1180–1192, Sep 2012.

[8] S. Long and Holder L. B. Graph-based Shape Shape Analysis for MRI Classification. *International Journal of Knowledge Discovery in Bioinformatics*, 2(2):19–33, 2011.

[9] H. Sakoe and S. Chiba. Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 26 (1):43–49, Feb 1978.

[10] A. Udomchaiporn, F. Coenen, M. García-Fiñana, and V. Sluming. 3-D MRI Brain Scan Feature Classification Using an Oct-Tree Representation. In *Advanced Data Mining and Applications (ADMA'13)*, pages 229–240. Springer, 2013.

[11] M. Vlachos, Z. Vagena, P. S. Yu, and V. Athitsos. Rotation Invariant Indexing of Shapes and Line Drawings. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, CIKM'05, pages 131–138, 2005.

# Brain Tumour Grading in Different MRI Protocols using SVM on Statistical Features

Mohammadreza Soltaninejad[1]
msoltaninejad@lincoln.ac.uk
Xujiong Ye[1]
xye@lincoln.ac.uk
Guang Yang [2]
gyang@sgul.ac.uk
Nigel Allinson[1]
nallinson11@gmail.com
Tryphon Lambrou[1]
tlambrou@lincoln.ac.uk

[1] School of Computer Science, University of Lincoln, UK
[2] St. George's, University of London, UK

**Abstract**

In this paper a feasibility study of brain MRI dataset classification, using ROIs which have been segmented either manually or through a superpixel based method in conjunction with statistical pattern recognition methods is presented. In our study, 471 extracted ROIs from 21 Brain MRI datasets are used, in order to establish which features distinguish better between three grading classes. Thirty-eight statistical measurements were collected from the ROIs. We found by using the Leave-One-Out method that the combination of the features from the 1st and 2nd order statistics, achieved high classification accuracy in pair-wise grading comparisons.

**Keywords**, Brain tumour grading, MRI images, superpixel segmentation, pattern recognition, SVM classification.

# 1 Introduction

Brain tumours are caused by abnormal and uncontrolled growing of the cells inside the brain or spinal canal. The primary tumours are those started in the brain and are categorised in four main types: Gliomas, Meningiomas, Pituitary adenomas and Nerve sheath tumours. The most popular grading system for tumours is that suggested by the World Health Organization (WHO). Regarding to the WHO grading system, the tumours are graded from I to IV, corresponding to least advanced to the most advanced diseases, respectively.

Utilizing computer-aided procedures for medical diagnosis and treatment is a growing field of research nowadays. Image processing and pattern recognition algorithms are widely used for analysis and interpretation of medical images. Feature extraction is the most important and impartible element of classification and pattern recognition tasks. In

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

the case of medical images, such as MRI, the reduction of dimensionality is of high importance. MRI images are three-dimensional volumetric data acquired with different protocols, which lead to extraction of high dimensional information in the form of statistical features. Classification of high dimensional data is based on these extracted features.

Several techniques have been used to detect and classify brain tumours in MRI images. Joshi et al. [1] developed a system for detection of Astrocytoma cancer tumours and classify them base on artificial neural network. They use grey-level co-occurrence matrix for texture feature selection and Neuro-Fuzzy classifier to classify the tumours. Georgiadis et al. [2] proposed a method to classify primary tumours and metastatic, which are originated outside brain. They applied non-linear least square feature transformation and combined it with a probabilistic neural network (PNN). Zacharaki et al. [3] used SVM with recursive feature elimination for classification of tumour grades in MRI images using texture, shape and rotation invariant texture features. They suggested one-versus-all SVM classification and majority voting for multi-class problem.

In this paper, we intend to classify tumour grades II, III and IV using different MRI acquisition protocols i.e. FLAIR, and T2. Region of interest segmentation is performed separately for each protocol. The segmentation is based on a superpixel based method with nearly similar intensity features. Then 1$^{st}$ order and 2$^{nd}$ order statistical textural features are extracted. The aim is to classify the tumours using the features collected from every individual protocol as well as their combinations. For this task, a support vector machine (SVM) classifier is utilized to classify different combinations of the data. The evaluation is performed using the overall classification accuracy and comparison of the results is based on the above mentioned protocols' features.

This paper is organised as follows. Section 2 explains the stages of the proposed method, which is partitioned into the main stages of the algorithm. Section 3 represents the experimental results. Finally, section 4 discusses the conclusion and future works.

# 2 Tools and Methods

## 2.1 Data Description

We acquired MR data using a GE Signa LX 1.5T MRI system (GE Healthcare, Milwaukee, WI, USA) equipped with a maximum field gradient strength of 22mT/m and using a quadrature head coil. The MRI sequences used in this study are described below:

FLAIR (TE = 133 ms, TR = 9000 ms, inversion time 2200 ms, band width = 61.04Hz). FLAIR were acquired in the axial plane with a field of view (FOV) 240 x 240 mm$^2$, matrix size 256 x 256 and 5 mm slice thickness with no slice gap.

Axial T2 weighted MRI using a single-shot spin echo echo-planar-imaging (EPI) sequence. The sequence covered whole brain with 50 contiguous slices that were acquired as two interleaved series of four repeats, i.e., 2.8 mm thick slices with 2.8 mm gaps. Other parameters were set as TR = 8000 ms, TE = 88 ms, acquisition matrix 96 x 96, FOV = 24 cm, and resulted in in-plane resolution of 2.5 mm.

Another T2 weighted MRI data were acquired using a dual echo sequence with FOV of either 220 x 220 mm$^2$ or 240 x 240 mm$^2$ , a 256 x 256 acquisition matrix, 29 slices with 5 mm thickness and a TR = 3500 ms and TE = 98 ms [4].

A cohort consisting of 21 patients entered retrospectively into our study, each with a brain tumour, who has been imaged with the aforementioned MRI sequences. The dataset consists of 6 grade II tumours, 4 grade III tumours, and 11 grade IV tumours. Each patient has a histological ground truth. Patient ages at the time of scanning ranged from 22 to 73 (mean 55), and consisted of 7 females and 14 males. All the MRI datasets have been realigned to remove eddy current effects by affine registration using FSL [5].

## 2.2 Superpixel Segmentation and Grouping

The image is segmented to small partitions using linear iterative clustering superpixel (SLIC) [6].The initial superpixel centres are located by uniform sampling the pixels in one slice in 2D space. The search space is a square window with the size of two times more than a superpixel size to decrease computations time. The cluster labels are updated in each iteration, based on the location and intensity distance of the pixels in search area to the centres. Therefore, the centres location is changed in every loop by calculating the gravity centres of new labels. A seed point is selected in an arbitrary slice by the user. The mean intensity value of the superpixel which contains this point is determined. The mean intensity value of the all the neighbouring superpixels are checked. The superpixels with values in a confidence margin of 3 are assumed to be in the region of interest. The process is repeated until we ensure that there is no new superpixel with similar value to add to the ROI.

## 2.3 Feature Extraction

Our statistical pattern recognition approach uses the classical steps of feature extraction, classification and feature selection, which are further described below.

The first step of our pattern recognition approach is the feature extraction step, which is the transformation of patterns into features that are regarded as a compacted representation. The usage of statistical features for the analysis and classification of textured images has been extensively demonstrated in the literature. Overall thirty eight statistical image features were collected from each image, given by category as:

The selected regions of interest (ROI) in the previous stage are based on the intensity mean of the superpixels. First Order Statistics [7] which express the distribution of grey levels (i.e. the intensity) within the selected ROI. These features are based on the histogram of the image, since they represent the frequency distribution of the grey level in the ROI. These 16 in total features are average, standard deviation, variance, mean of the absolute deviation, median absolute deviation, coefficient of variance, skewness, kurtosis, maximum , minimum, median and mode of the intensity values, central moments, range, interquartile range and entropy.

Second Order Statistics [7] which are measurements collected by the Grey-Tone Spatial-Dependent Matrices. These matrices of grey-tone spatial-dependence frequencies are a function of the angular relationship between neighbouring image elements, and additionally a function of the distance between them. This is performed for one neighbouring pixel in the selected ROI in four different directions [0°, 45°, 90°, and 135°]. However, in our case the average feature vector of the four directional feature vectors was estimated in order to reduce the dimensionality of the feature combinations. The features which can be extracted by the grey-tone spatial-dependence matrices are: autocorrelation, contrast, correlation 1, correlation 2, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of squares' variance, sum

average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation 1, information measure of correlation 2, inverse difference, and inverse difference moment normalized.

### 2.4 Support Vector Machine Classification

Support vector machine (SVM) is a supervised learning method, which is used for classification and regression tasks [8]. The data must include exactly two number of classes. The SVM finds the best hyperplane for separation of the classes, which presents the largest margin between them. The margin is defined as the maximum distance in which in the best case scenario the data points of the different classes are separable in the feature space. The aim of using SVM is to find a hyperplane in the feature space to separate them with the minimal error (i.e. maximum distance from the clusters closets points).

## 3 Results

The aim of this study is to examine the performance of the linear SVM classification on brain MRI datasets which have been delineated either manually or using a superpixel based method, and in particular, to determine whether we can distinguish between different grades of cancer brain tissue.

The performance of the linear SVM classifier was evaluated by using the Leave-One-Out method. In addition, for each set of features all possible combinations were tested up to three dimensional decision spaces. Those features, which achieve the best classification rate, were used in the pattern recognition process. This phase is called feature selection, and aims to reduce the features set to a subset, which consists only of meaningful information (i.e. features that characterize best) about the images we want to classify. The classification accuracy results presented in this paper are those, which fulfil the requirement that the overall accuracy is more than 80%.

Figure 1, shows three examples of grades II, III, and IV, acquired using the FLAIR MRI protocol. Each column presents the original image, as well as the results of the superpixel and the manual segmentation.

In terms of classification accuracy, we tested all the pair-wise combinations between the different grades present in our cohort. In addition comparisons between datasets using different acquisition MRI protocols were also performed. These results are presented in Tables 1 and 2, for ROIs extracted manual and with superpixel segmentation, respectively. It is worth noticing that the best classification results were achieved in three-dimensional space, for both manual and superpixel ROIs.

In terms of the performance of the Statistical Features extracted from all the brain MRI images, we concluded that: Features from the 1$^{st}$ Order Statistics obtained from either the manual or the superpixel segmentation, produced classification accuracy results above the thresholds set. Features from the 2$^{nd}$ Order Statistics obtained from either the manual or the superpixel segmentation produced the best classification accuracy results. The overall number of feature combinations which exceeded the inclusion threshold of 80.00%, is greater on features from the superpixel segmentation than that of the manual delineation (34652 to 29429).

The usage of statistical features for the analysis and classification of textured images has been extensively demonstrated in the literature. Our results suggest that features from

the 2nd Order Statistics achieved the best classification accuracy results, since such measurements focus on the overall nature of the texture such as homogeneity, contrast, the presence of organized structure, complexity, and the grey tone transitions within the image.

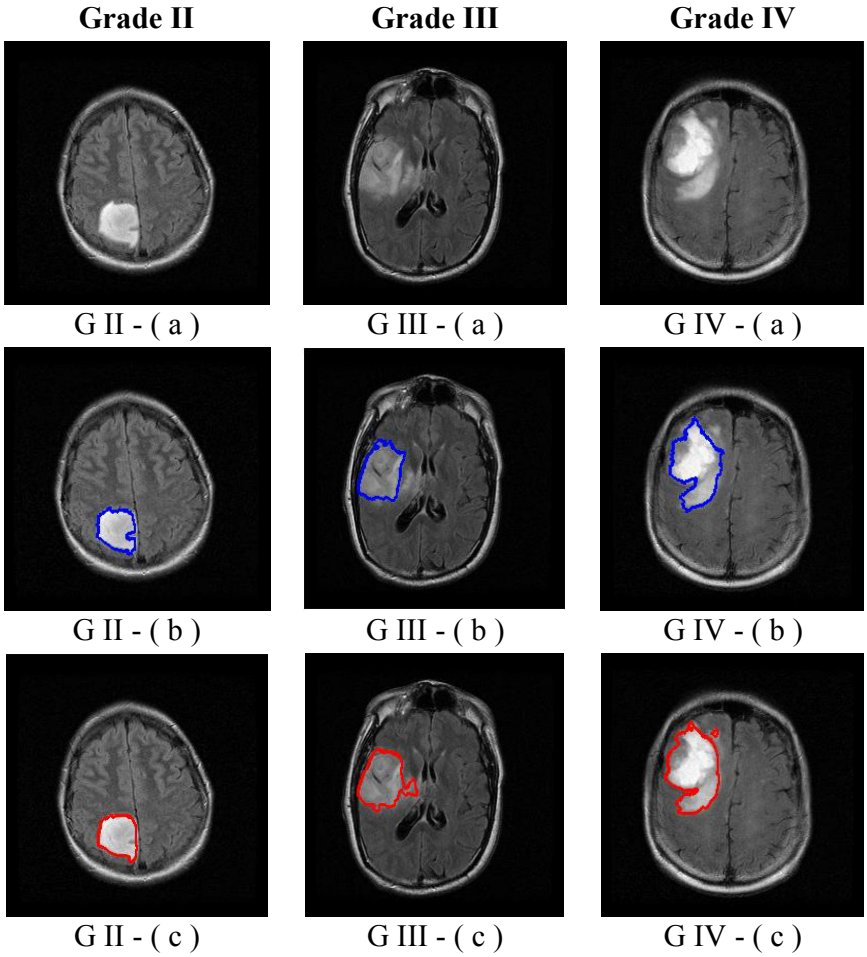



Figure 1: Examples of Grade II, III, and IV brain oedemas: a) original MRI slice, b) superpixel segmentation, c) manual segmentation.

| MANUAL SEGMENTATION | | | | | | |
|---|---|---|---|---|---|---|
| | 2D | | | 3D | | |
| **Grade Combinations** | **FLAIR Accuracy** | **T2 (1) Accuracy** | **T2 (2) Accuracy** | **FLAIR Accuracy** | **T2 (1) Accuracy** | **T2 (2) Accuracy** |
| **II vs III** | 90.00 | 80.00 | 90.00 | 90.00 | 80.00 | 90.00 |
| **II vs IV** | 94.12 | 92.36 | 94.12 | 100.00 | 82.36 | 94.12 |
| **III vs IV** | 93.33 | 93.33 | 86.67 | 93.33 | 93.33 | 93.33 |
| **II+IV vs III** | 90.48 | 85.71 | 85.71 | 90.48 | 85.71 | 90.48 |
| **II+III vs IV** | 90.48 | 80.95 | 90.48 | 95.24 | 80.95 | 90.48 |
| **III+IV vs II** | 90.48 | 80.95 | 90.48 | 95.24 | 80.95 | 90.48 |

Table 1: Overall classification accuracy for 2D and 3D decision space, using the manual segmented ROIs.

| SUPERPIXEL SEGMENTATION | | | | | | |
|---|---|---|---|---|---|---|
| | 2D | | | 3D | | |
| Grade Combinations | FLAIR Accuracy | T2 (1) Accuracy | T2 (2) Accuracy | FLAIR Accuracy | T2 (1) Accuracy | T2 (2) Accuracy |
| II vs III | 100.00 | 80.00 | 80.00 | 100.00 | 80.00 | 100.00 |
| II vs IV | 88.24 | 88.24 | 94.12 | 94.12 | 94.12 | 100.00 |
| III vs IV | 86.67 | 93.33 | 86.67 | 93.33 | 93.33 | 93.33 |
| II+IV vs III | 90.48 | 90.48 | 85.71 | 95.24 | 95.24 | 90.48 |
| II+III vs IV | 90.48 | 80.95 | 85.71 | 90.48 | 85.71 | 85.71 |
| III+IV vs II | 90.48 | 85.71 | 85.71 | 90.48 | 90.48 | 85.71 |

Table 2: Overall classification accuracy for 2D and 3D decision space, using the superpixel segmented ROIs.

# 4 Conclusion

In this paper a feasibility study of brain MRI dataset classification, using ROIs which have been segmented either manually or using a superpixel based method in conjunction with statistical pattern recognition methods is presented. In our study, 471 extracted ROIs from 21 Brain MRI dataset were used, in order to establish which features distinguish better between three grading classes. Thirty-eight statistical measurements were collected from the ROIs. We found by using the Leave-One-Out method that the combination of the features from the 1st and 2nd Order statistics, achieved overall classification accuracy above the self-imposed threshold of 80.00% that we had originally set.

# References

[1] D. M. Joshi, N. K. Rana, V. M. Misra. Classification of brain cancer using artificial neural network, Proceedings of the 2010 2nd International Conference on Electronic Computer Technology, 112-116, Kuala Lumpur, Malaysia, May 2010.

[2] P. Georgiadis, D. Cavouras, I. Kalatzis, A. Daskalakis, G. C. Kagadis, K. Sifaki, M. Malamas, G. Nikiforidis, E. Solomou, Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features, Computer Methods and Programs in Biomedicine, 89 (1) : 24-32, 2008.

[3] E. I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E. R. Melhem, C. Davatzikos. Classification of brain tumour type and grade using MRI texture and shape in a machine learning scheme, Magnetic Resonance in Medicine, 62(6): 1609-1618, 2009.

[4] T. Jones, B. Bell, and T. Barrick, "A novel whole-brain DTI segmentation technique for brain tumour delineation and diagnosis," in Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM), 2012, vol. 20, p. 188.

[5] M. Jenkinson, C.F. Beckmann, T.E. Behrens, M. W. Woolrich, and S. M. Smith. Fsl. NeuroImage, 62:782–790, 2012.

[6] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34 (11): 2274-2282, 2012.

[7] A.K. Jain, "Fundamentals of Digital Image Processing", Prentice-Hall International Editions, N.J., USA, 1989.

[8] N. Cristianini, and J. Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, First Edition, Cambridge: Cambridge University Press, 2000.

# New Optical Mapping Design and Automated Algorithms for Cardiac Electrophysiology

Ting Yue Yu[1]
tyy608@bham.ac.uk
Hamid Dehghani[2]
Keith Brain[3]
Fahima Syeda[4]
Andrew Holmes[4]
Paulus Kirchhof[4]
Larissa Fabritz[4]

[1] PSIBS Doctoral Training Centre
[2] School of Computer Science
College of Engineering and Physical Sciences
[3] Pharmacology and Therapeutics
School of Clinical and Experimental Medicine
[4] Centre for cardiovascular Sciences
School of Clinical and Experimental Medicine
College of Medical and Dental Sciences

University of Birmingham
Edgbaston, B15 2TT, UK

**Abstract**

Optical mapping is an important tool for assessment of cardiac electrophysiology. We demonstrate a new system for quantification and measurement of electrophysiological parameters in isolated cardiac tissue. The system makes use of voltage sensitive fluorescent dyes that shift in spectral property in response to millivolt changes in potential across cell membranes. Automated analysis of the pixel-wise measurements yields information on action potential durations and isochronal maps allowing for high throughput of data analysis. The algorithms that we propose reliably describe activation sequences and allow for quantification of conduction velocities.

## 1 Introduction

Cardiovascular disease is the largest cause of death worldwide and atrial fibrillation (AF) is the most common arrhythmia. AF regularly causes strokes and cardiac death. Despite the progress made on the characterisation of the factors that lead to AF there is still a need to further understand disease mechanisms to improve therapies for prevention and management [2]. Genetic mouse models pose as an attractive tool for these studies [7] but their small size particularly the atria prove challenging for detailed assessment. Despite this, they allow for characterisation of functional molecular consequences of their genetic alterations.

Optical mapping with its high temporal and spatial resolution is a valuable technique in aiding our understanding of arrhythmias[1]. This method utilises voltage sensitive dyes;

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

these bind on to the lipid molecules on a cell membrane and will shift its spectral properties in response to a change in potential[6]. The motivation for development of these dyes arise where microelectrode measurements are unsuitable, and makes possible the measurement of spatial and temporal variations in membrane potential, hence it is it ideal for cardiac tissue[6]. High speed and high quantum efficient cameras are particularly well suited for imaging cardiac activation as events occur over the millisecond time scale and changes in fluorescence intensity are small. This is important, in particular for small animal preparations as heart rates are a magnitude faster than humans. These studies at high acquisition rates involve generating large data sets, in the range of hundreds of thousands of images, and hence the need for faster analysis methods. The use of semi-automated algorithms not only allows for signal processing but eliminates any user bias saving time and increasing reliability. One difficulty that arises from imaging cardiac muscle is its contractile properties causing movement artefacts; these greatly affect the emitted fluorescence but can be overcome with the use of mechanical uncouplers in the experimental setting which inhibit these muscle contractions.

There are several imaging techniques for the study of the heart such as computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound [3]. These are mostly limited to visualising diameters and contractile function but not electrical activation and repolarisation. Another method in which activation spread can be investigated is to use contact mapping [4] where an array of electrodes are placed on the surface, but these have limited use with the resolution determined by the number of electrodes that can be physically placed on the sample and there is no information on repolarisation.

In cardiac tissue, as the cells depolarise, they cause a sharp upstroke in membrane voltage. This is followed by a slower rate of repolarisation. In order to understand cardiac arrhythmias from optical mapping data we analyse the duration of these action potentials at 30, 50 and 70% of the repolarisation. These action potential durations (APD) are important in determining the membrane voltage as this is characteristic of ion channel activity. The different phases of an action potential are caused by ion channel currents. In addition to this, it is useful to image activation patterns from a sample by generating an isochronal map. This allows for determining spatial areas of slow or fast conduction, which is particularly important in evaluating disease models. In this paper we discuss the development of the optical mapping design along with the algorithms to assess the images produced.

# 2 Method

## 2.1 Optical mapping experimentation and design

Isolated murine hearts were perfused with the potentiometric dye Di-4-Anepps, with a peak Ex/Em of 502/703nm. This dye is fast responding and has been used in a multitude of other preparations not only murine hearts, but in larger animals and cells such as neurons[6]. After 5 minutes the atria were individually dissected and superfused with oxygenated solution and the excitation-contraction uncoupler blebbistatin. Preparations were continuously superfused and stimulated at twice the diastolic threshold with a 2ms pulse. The stimulation frequencies ranged from 3.33 Hz to 12.5 Hz. Samples were field illuminated by two twin 530nm LEDs and the emitted light was filtered at 630nm. Images were acquired at up to 2 kHz on the CMOS camera using WinFluor V3.4.9 (Dr John Dempster, University of Strathclyde, UK). From the raw images, time course information from regions of interest

was extracted into text format for APD measurements or an image series to tiff images for isochronal mapping.

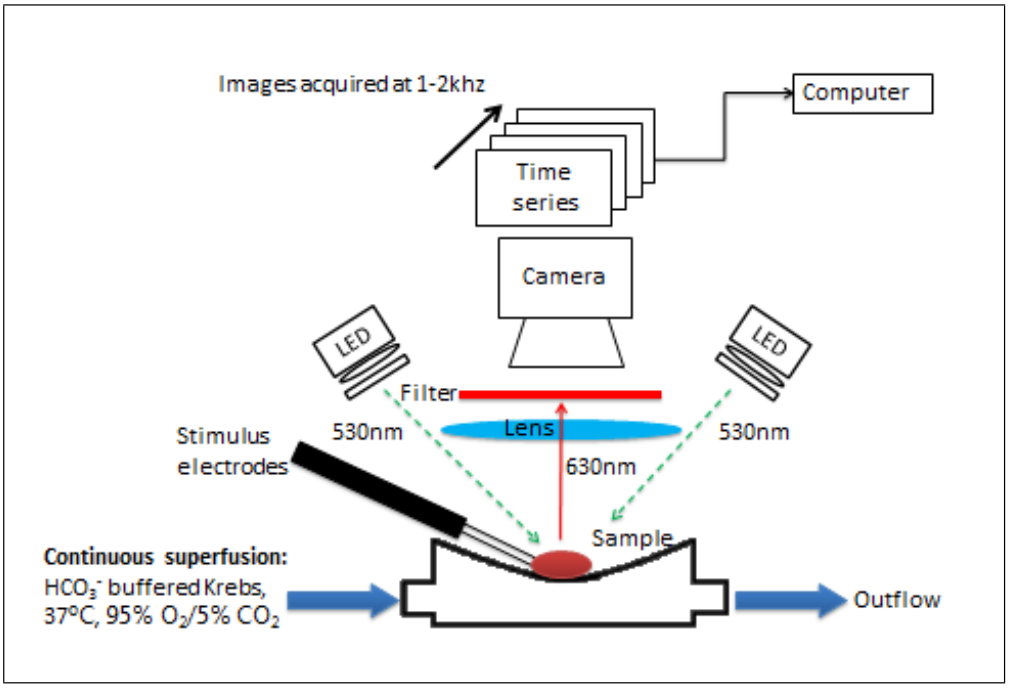



Figure 1: Optical imaging system. Samples were continuously superfused and stimulus electrodes were used to pace the heart at desired frequencies. Four LEDs were required to provide an even field illumination of the sample.Images acquired on high speed CMOS camera. (Figure adapted from submission to *Progress in Biophysics and Molecular Biology*, under review)

## 2.2 Optical action potential recordings

The fluorescence intensity obtained from the raw image data is inversely proportional to membrane voltage and the action potential properties can be measured pixel-wise from this time varying signal. The obtained optical traces often exhibit large shifts in baseline which is sometimes attributed to photobleaching and variations in solution level since the shifts do not follow a predetermined pattern. A linear top hat filter with a suitable structuring element was appropriate for flattening the signal baseline without affecting the action potential morphology. With high speed acquisition, noise is a strong contributing factor in image formation. To reduce noise, up to 25 action potentials were grouped together and averaged, but firstly each action potential was identified by using the 'findpeaks' detection algorithm in MATLAB. The next step was to identify the start of the action potential. This is conventionally chosen to be the fastest upstroke of the depolarisation. This was simply calculated by taking the peak of the signal derivative. The mean value of ten frames before the upstroke determined the baseline. This, along with the peak height, yields the absolute signal amplitude. From here it was possible to calculate the APDs at 30, 50, 70% repolarisation by linearly interpolating the data.

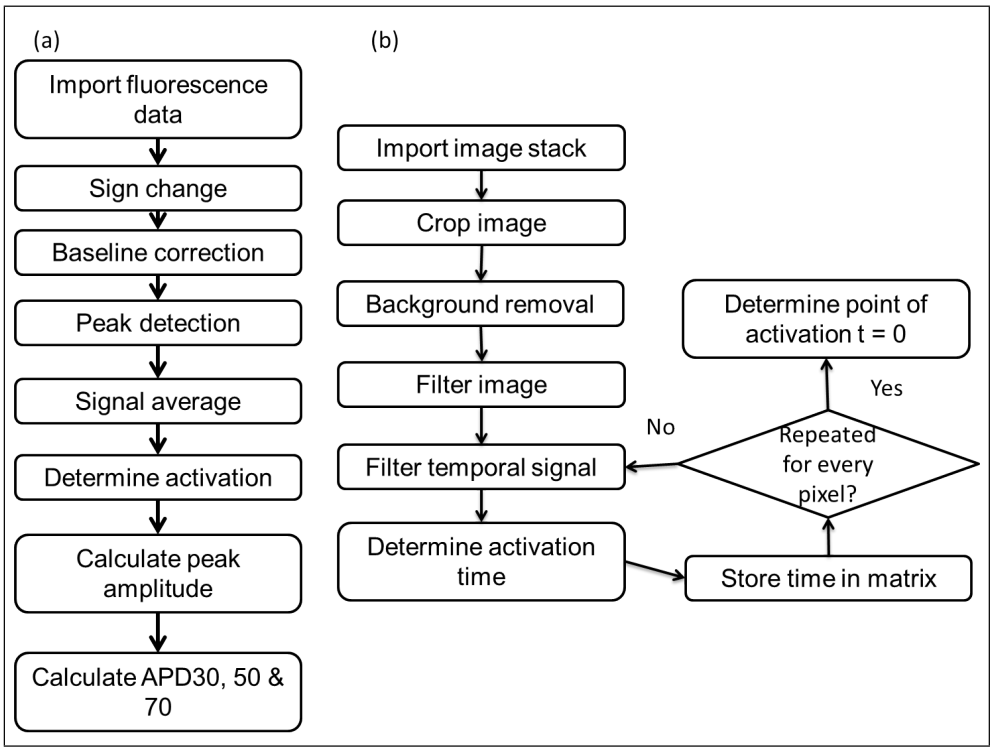



Figure 2: Outline of steps used to determine action potential durations (a) and generation of isochronal maps (b)

## 2.3 Isochronal maps

To create a map showing the different regions of activation the background must be first removed as this interferes with the subsequent processes. As the fluorescence signal tends to always be greater than the background, thresholding proved to be the most suitable method. Following this, a mean filter was applied to smooth the images. The next step was to determine the point of activation for every pixel. This was a similar process to that described in optical action potential recordings above, but signals were not averaged or did not have their durations measured. However, a Savitzky-Golay filter was applied to each pixel in the time series smoothing the data. This step was necessary to reduce noise amplification during signal differentiation. The peak of this derivative was chosen to be the activation time, repeating this process for every pixel generates the isochronal map. The origin of activation (t=0) was automatically determined by arranging the unique values of this array in ascending order. Each of these values represents a time determined by the exposure of the camera, and the second differential of these values should yield close to zero. The longest sequence of zeroes indicated the start point of activation spread. As each pixel represents a time value, the distance between two points of activation was used to represent a conduction velocity.

# 3 Results

An optical mapping system was built and used on small cardiac tissue samples. The automated algorithms described above accurately determined APD values from signal traces. The quality of these images also allowed for the generation of isochronal maps, which indicate

the changes in activation across a tissue. From Figure 3(a) the baseline is well defined, this is subtracted from the signal and allows for analysis of action potentials. From Figure 3(b) we can see that the averaging can accurately combine several action potentials as indicated by the small errors. This allows for APD measurements as indicated by Figure 3(c). Validation work has been completed using microelectrode and monophasic action potential (MAP) and is currently under review, but results remain consistent throughout all techniques. These traces are derived from the temporal data acquired from a selected region of interest typically a 4x4 area, an example of a raw image can be seen in Figure 4.

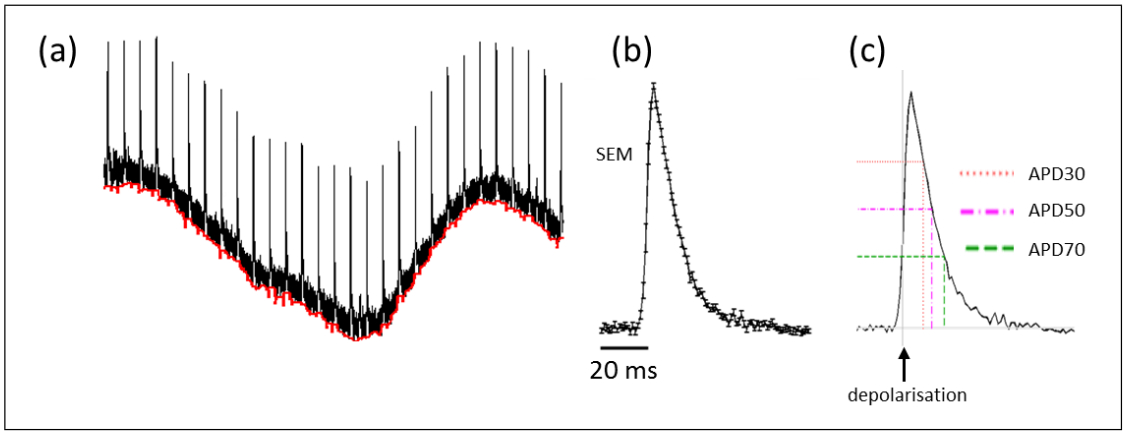



Figure 3: (a) An example of a drifting signal, the red line indicates the baseline detected. (b) An average of 10 action potentials: error bars indicates standard errors of the mean. (c) Displays the calculation of APD values and the point of activation indicated by the black arrow. (Figure (b) and (c) adapted from submission to *Progress in Biophysics and Molecular Biology*, under review)

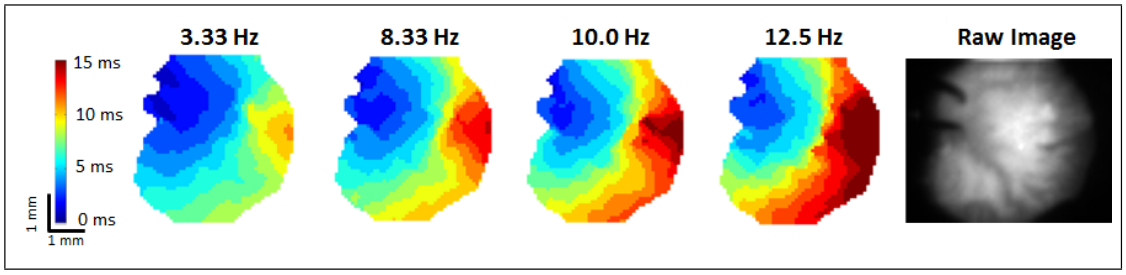



Figure 4: Isochronal maps showing activation patterns across an isolated left atrium and the raw image(left). These high resolution maps indicate that at higher stimulation frequencies the activation time increases across the entire tissue.

## 4 Discussion and Conclusion

This paper has presented a new optical mapping system capable of imaging isolated atria. The algorithms used have also been shown to perform APD calculations autonomously from fluorescence data, increasing signal quality by use of averaging. In addition to this, the baseline correction was proven to be robust and efficient. Other methods such as polynomial

fitting have limited use with signals that fluctuate greatly [5]. The isochrones were generated automatically, the only input was the user cropping the image to highlight the region of interest. As seen in Figure 4, the algorithms used to determine t=0, the activation start point, can be automated saving time, allowing for faster analysis. This differs from most current methods where manual selection is required, which can introduce subjectivity to results. This system was also sensitive enough to detect changes in activation at different stimulus frequencies. With these accurate measurements this system can be used in studies for characterising arrhythmias in mouse models. This is important in enhancing knowledge of treatments for those afflicted with cardiovascular disease.

# References

[1] A Blana, S Kaese, L Fortmuller, S Laakmann, D Damke, K van Bragt, J Eckstein, I Piccini, U Kirchhefer, S Nattel, G Breithardt, P Carmeliet, E Carmeliet, U Schotten, S Verheule, P Kirchhof, and L Fabritz.

[2] A John Camm, Sana M Al-Khatib, Hugh Calkins, Jonathan L Halperin, Paulus Kirchhof, Gregory Y H Lip, Stanley Nattel, Jeremy Ruskin, Amitava Banerjee, Dan Blendea, Eduard Guasch, Matthew Needleman, Irina Savelieva, Juan Viles-Gonzalez, and Eric S Williams. A proposal for new clinical concepts in the management of atrial fibrillation. *American heart journal*, (3):292–302.e1, September . ISSN 1097-6744. doi: 10.1016/j.ahj.2012.05.017.

[3] L Fabritz, L Fortmuller, T Y Yu, M Paul, and P Kirchhof. Can preload-reducing therapy prevent disease progression in arrhythmogenic right ventricular cardiomyopathy? Experimental evidence and concept for a clinical trial. *Prog Biophys Mol Biol*, 110(2-3): 340–346, 2012. doi: 10.1016/j.pbiomolbio.2012.08.010.

[4] P Kirchhof, E Marijon, L Fabritz, N Li, W Wang, T Wang, K Schulte, J Hanstein, J S Schulte, M Vogel, N Mougenot, S Laakmann, L Fortmueller, J Eckstein, S Verheule, S Kaese, A Staab, S Grote-Wessels, U Schotten, G Moubarak, X H Wehrens, W Schmitz, S Hatem, and F U Muller. Overexpression of cAMP-response element modulator causes abnormal growth and development of the atrial myocardium resulting in a substrate for sustained atrial fibrillation in mice. *Int J Cardiol*. doi: 10.1016/j.ijcard.2011.10.057.

[5] Jacob I Laughner, Fu Siong Ng, Matthew S Sulkin, R Martin Arthur, and Igor R Efimov. Processing and analysis of cardiac optical mapping data obtained with potentiometric dyes. *American journal of physiology. Heart and circulatory physiology*, (7):H753–65, October . ISSN 1522-1539. doi: 10.1152/ajpheart.00404.2012.

[6] L M Loew. Potentiometric dyes: Imaging electrical activity of cell membranes. *Pure and Applied Chemistry*, 68(7):1405–1409, 1996. doi: DOI10.1351/pac199668071405.

[7] G Riley, F Syeda, P Kirchhof, and L Fabritz. An introduction to murine models of atrial fibrillation. *Front Physiol*, page 296. doi: 10.3389/fphys.2012.00296.

# Multiscale AM-FM Models and Instantaneous Amplitude Evaluation for Mammographic Density Classification

Ioannis Constantinou[1]
ikonst01@cs.ucy.ac.cy
Marios Pattichis[2]
pattichis@gmail.com
Chrysa Tziakouri[3]
kithreas@cytanet.com.cy
Constantinos Pattichis[1]
pattichi@ucy.ac.cy
Styliani Petroudi[1]
styliani@ucy.ac.cy

[1] Department of Computer Science
University of Cyprus
Nicosia, Cyprus

[2] Department of Electrical and Computer Engineering
University of New Mexico
New Mexico, USA

[3] Nicosia General Hospital
Nicosia, Cyprus

**Abstract**

Breast cancer is the most common cancer in women and the number of incidences keeps rising. Mammographic breast density has been recognized as a very important breast cancer risk and can also mask abnormalities. Information regarding mammographic breast density may be used for planning individualized breast cancer screening and treatment. Thus, breast density is increasingly assessed as the limitations of a one-fits-all method of screening and Computer Aided Detection become more apparent. The presented work investigates the use of Amplitude-Modulation Frequency-Modulation (AM-FM) models in the evaluation of multiscale Instantaneous Amplitude (IA) features for the characterization of breast density. AM-FM models provide a meaningful and concise method to model digital images. The IA evaluated at different frequency scales is used to capture the relative variations in the breast tissue characteristic to the different breast density classes. Normalized histograms of the IA across the different frequency scales - estimated using multiscale Dominant Component Analysis - are used to model the breast density classes. Classification of a new mammogram into one of the density categories is achieved using the k-nearest neighbor method and the Euclidean distance metric. The method is evaluated using the Breast Imaging Reporting and Data System on the Medical Image Analysis Society mammographic database and the results are presented and compared to other methods in the literature. The presented method allows breast density classification accuracy reaching over 80%.

## 1 Introduction

Breast cancer is the second most commonly diagnosed cancer, and the most common cancer in women both in the developing and developed world. Mammography has been the modality of choice for breast cancer screening for early detection. Yet, the incidences of breast

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

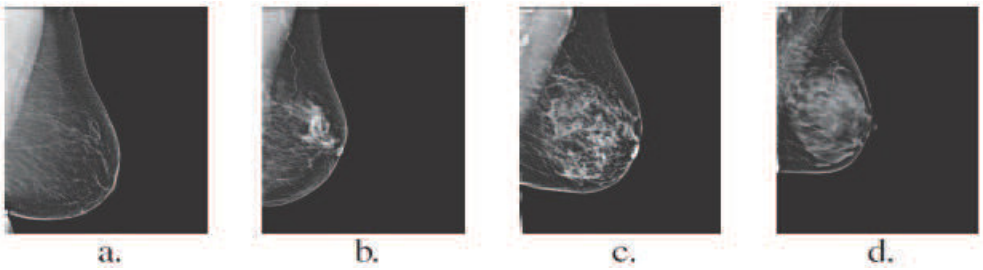



Figure 1: Examples of mammograms from the breast parenchymal density BI-RADS classes a) BI-RADS I, b) BI-RADS II, c) BI-RADS III, d) BI-RADS IV.

cancer continue to rise. Breast density has not only been shown to be one of the most important risks for developing breast cancer [3], but it also impacts the ability of the detection of breast cancer by masking abnormalities. Breast density assessment information can be used towards breast cancer screening personalization; as the basis for supplementary screening using other imaging techniques [11]. Mammographic breast density refers to the amount of fibroglandular tissue in the breast as it appears on a mammogram, first reported by Wolfe in 1976 [13]. The Breast Imaging Reporting and Data System (BI-RADS) Atlas produced by the American College of Radiology (ACR) [1] provides a mammographic breast density categorization that serves as a comprehensive guide and a classification system, important for peer review and quality assurance. The BI-RADS mammographic breast density categories descriptions are [1]: (I) the breast is almost entirely fatty, (II) there are scattered fibroglandular densities, (III) the breast is heterogeneously dense which may obscure small masses, and (IV) the breast is extremely dense, which lowers the sensitivity of mammography. The first two classes are low density, low risk and the last two are high density, higher risk. Examples of the four BI-RADS mammographic breast density classes can be seen in Fig. 1.

The new requirements regarding reporting breast density in place, the global increase in breast cancer and the high inter- and intra- observer variability in breast density classification [6] necessitate the development of accurate mammographic breast density classification algorithms. A review can be found in [9]. Petroudi et al. [10] proposed a scheme that uses texture models to capture the mammographic appearance: parenchymal density patterns are modeled as a statistical distribution of clustered, rotationally invariant filter responses in a low dimensional space. More recently He et al. [5] calculated different texture features on the intensity histograms to characterize different tissue types, using Bayes classifier for classification.

Amplitude-Modulation Frequency-Modulation (AM-FM) methods provide powerful and physically meaningful image decompositions that describe non-stationary content and capture local (instantaneous) variations in amplitude, frequency, and phase. AM-FM models characterize each pixel value as sampled from amplitude and frequency modulated orthogonally oriented signals. The AM components capture the local texture contrast whilst the FM components capture the local texture orientation and granularity. The extraction of such features is most relevant to the appearance of fibroglandular density patterns and thus AM-FM derived features may provide a strong basis for the development of density classification algorithms. AM-FM models have been used in different medical image applications and a recent review can be found in [8]. This paper presents an initial evaluation of the use of AM-FM features for breast density classification towards the development of a complete

Figure 2: The Gabor filterbank used.

Computer Aided Detection (CAD) system incorporating breast density. Multi-scale Instantaneous Amplitude (IA) AM-FM features are used to model mammographic density and classification is achieved using k-nearest neighbor (k-NN) with Euclidean distance.

## 2 Method

Initially the mammograms are pre-processed and normalized, and the breast region is segmented using the methodology in [10]. The AM-FM demodulation is evaluated on the pre-processed mammograms but only the values corresponding to the breast region mask are utilized for tissue characterization. Each segmented breast region is then represented by a set of normalized histograms of the evaluated AM-FM features, one for each frequency scale. Finally, the k-NN method is used to classify mammograms to the corresponding density class with the leave one woman out cross-validation method.

AM-FM methods model an image as a function of spatially local amplitude and frequency modulations. An image can be decomposed to a sum of AM-FM components using:

$$I(x,y) = \sum_{n=1}^{n=M} a_n(x,y)\cos(\varphi_n(x,y)). \tag{1}$$

In the AM-FM expansion $I(.)$ is the input image, $a_n(x,y)\cos(n\varphi(x,y))$ is the collection of $M$ AM-FM component signals used to model the essential image modulation structure [7], and $n = 1,2,...,M$ denote the different frequency scales. Each scale is defined in terms of a collection of bandpass filters that share similar magnitude range. $a_n(x,y)$ denote the instantaneous amplitude (IA) functions, and $\varphi_n(x,y)$ denote the instantaneous phase functions (IP). The IA amplitude reflects local image intensity variations, e.g. edges, with different spatial scale variations reflected in different frequency scales. There are different AM-FM demodulation methods that can be used for the estimation of the different AM-FM components. The multiscale approach introduced by Murray et al. [7] is used here. Very briefly, given the real input image $I(x,y)$ the 2-D extended analytic signal associated with $I(x,y)$ is computed by:

$$I_{AS}(x,y) = I(x,y) + jH_{2d}[I(x,y)]. \tag{2}$$

where $H_{2d}$ denotes the $2-D$ extension of the $1-D$ Hilbert transform operator [7]. The resulting $I_{AS}$ is processed through a filterbank. For each bandpass filter output $I_{AS_n}$, it is possible to estimate the IA with:

$$a_n(x,y) \approx |I_{AS_n}(x,y)| \tag{3}$$

Unlike [7] the Gabor filterbank using eight orientations and six different frequency scales is applied [2] (see Fig. 2). Image characterization is achieved using multiscale Dominant Component Analysis (mDCA). The filterbank channels are grouped into scales from very

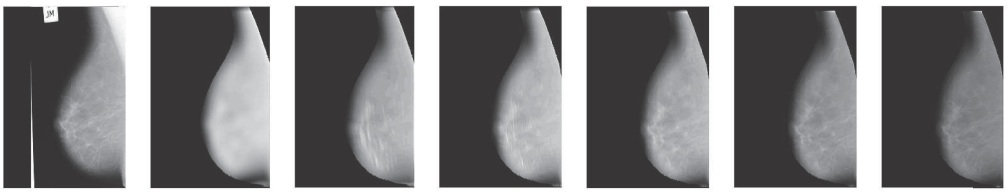


Figure 3: The maximum IA representation for all scales ranging from the very very low frequencies to the very high for a BI-RADS1 mammogram. The first image is the original mammogram.

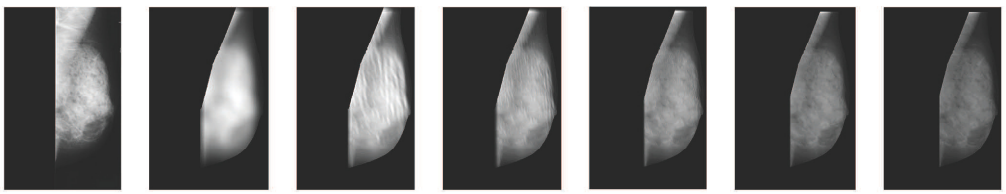


Figure 4: The maximum IA representation for all scales for a BI-RADS4 mammogram.

very low frequencies to very high frequencies [2] and the maximum IA is evaluated across all orientations for each scale, and used to characterize each pixel in the breast region. Examples can be seen in figures 3 and 4. For each scale, for each mammogram, the normalized histogram of the IA is evaluated. Normalized histograms from different frequency scales are concatenated and used to characterize the corresponding mammogram and in turn the different mammographic density classes.

The k-NN method is used to classify mammograms to the corresponding density class using the Euclidean distance measure. $k$ is set to 5 after empirical evaluation. Despite k-NN being a quite simple classification method - for this stage of the algorithm development it works quite well. The method is developed and quantitatively evaluated using the 206 mammograms with no abnormalities present in the Medical Image Analysis Society (MIAS) mammogram database [12]. The mammograms are classified into one of the four BI-RADS density classes - as classified for [9]. The performance of the presented method is evaluated using the leave one woman out validation model on the normalized histograms of the presented AM-FM texture features.

## 3 Results - Discussion

The algorithm is evaluated on the mammograms that do not include any abnormalities from the MIAS mammographic database [12]. The classification accuracy using k-NN on the different density classes is shown in Table 1. The agreement with the density annotations provided with the MIAS database, when all the mammograms from all different classes are used for classification, is 80.00%. Accuracy is calculated as the percentage of correctly classified mammograms in a breast parenchymal density category over the ground truth total number of mammograms in that category.

The presented work has demonstrated that AM-FM features can help differentiate be-

Table 1: Classification accuracy results for the MIAS breast density characterization using a k-NN classifier with IA AM-FM features

| MIAS BI-RADS Class | BI-RADS 1 | BI-RADS 2 | BI-RADS 3 | BI-RADS 4 | TOTAL |
|---|---|---|---|---|---|
| IA AM-FM | 89% | 85% | 69% | 77% | 80% |

tween the different breast density classes. So far, only histograms of the IA have been used which capture the contrast in the texture of the images. In addition, the value of IA reflects the extend of the presence of frequency components from that particular scale in the image. Figures 3 and 4 provide examples of the IA for the 6 scales and it can be seen that the AM-FM features can capture differences in the mammographic appearance of fibroglandular tissue e.g homogeneity vs heterogeneity that characterize different mammographic density variations or lack thereof. This was also investigated by classifying the images using the IA histogram from each frequency scale separately, which resulted in different classification accuracies for the different classes e.g. the very low frequency scale captured the homogeneity present in very high density breasts providing the best single scale classification accuracy for BI-RADS 4 at 71%. Depending on the application, different AM-FM decompositions using different frequency coverage can provide for higher classification accuracy, and so can the incorporation of IF and IP. Moreover, the use of all concatenated values may negatively affect the results because it can affect the evaluation of the measure distance. The next step will be dimensionality reduction using Principal Component Analysis and other dimensionality reduction methods. The presented methodology did not involve any processing or reduction of the IA features. Still, the resulting classification accuracy of 80% into one of the four BI-RADS classes compares very favorably with other methods in the literature. Petroudi et al. [10] evaluated texton histograms using chi-square distance and achieved a classification accuracy of 76% but on a different database. Oliver et al. [9] extracted morphological and texture features from the segmented breast areas and used a Bayesian combination of a number of classifiers, achieving 84% BI-RADS classification accuracy on the same dataset. Chen et al. [4] achieved a classification density of 76% on the MIAS database with BI-RADS classification using a topographic representation, saliency and shape. The achieved classification accuracy from the evaluation of the IA warrants further investigation of AM-FM texture features to help establish a best feature selection for each density class.

# 4 Conclusion

The evaluation of the IA from AM-FM provides a characterization of the contrast of texture at different frequency scales for an image. Thus, it may be used to capture different characteristics in the appearance of mammographic breast density for BI-RADS density classification. The method builds on the AM-FM demodulation method presented in [7] but with the use of a Gabor filterbank. The normalized histogram of the maximum IA across all frequencies for each scale is used to characterize the breast density for each mammogram. The histograms are concatenated to characterize each image and classification is achieved by comparing the corresponding distribution to the rest using k-NN and Euclidean distance. The achieved classification accuracy of 80% on the MIAS [12] database for BI-RADS classification is high

and comparable to other methods in the literature. Future work will involve use of additional AM-FM features, dimensionality reduction and investigation of other classification methods.

# Acknowledgements

This work is supported by the Cyprus Research Promotion Foundation's Grant ΤΠΕ/OPIZO/0311(BIE)/029 and is co-funded by the Republic of Cyprus and the European Regional Development Funds.

The authors would like to thank J. Marti from the University of Girona and J.Pont and E.Pérez from the Department of Radiology of the Girona University Hospital Dr. Josep Trueta (Spain) and E.R.E. Denton from the Department of Breast Imaging of the Norwich and Norfolk University Hospital (UK) for the BIRADS classification of the MIAS database used in this work.

# References

[1] ACR American College of Radiology. *ACR BIRADS Atlas Breast Imaging Reporting and Data System*. American College of Radiology, 5th edition, 2013.

[2] A.C. Bovik, N. Gopal, T. Emmoth, and A. Restrepo. Localized measurement of emergent image frequencies by gabor wavelets. *IEEE Trans Inf Th*, 38(2):691–712, 1992.

[3] N.F. Boyd, J.W. Byng, R.A. Jong, E.K. Fishell, L.E. Little, A.B. Miller, G.A. Lockwood, D.L. Tritchler, and M.J. Yaffe. Quantitative classification of mammographic densities and breast cancer risk: Results from the canadian national breast screening study. *J Natl Cancer Inst*, 87(9):670–675, 1995.

[4] Z. Chen, A. Oliver, E. Denton, and R. Zwiggelaar. Automated mammographic risk classification based on breast density estimation. In *Pat Rec and Im Anal*, volume 7887 of *LNCS*, pages 237–244. Springer, 2013.

[5] W. He, E. Denton, and R. Zwiggelaar. Mammographic segmentation and risk classification using a novel binary model based bayes classifier. In *Breast Imaging*, volume 7361 of *LNCS*, pages 40–47. Springer, 2012.

[6] M.B.I. Lobbes, J.P.M. Cleutjens, V. Lima Passos, C. Frotscher, M.J. Lahaye, K.B.M.I. Keymeulen, R.G. Beets-Tan, J. Wildberger, and C. Boetes. Density is in the eye of the beholder: visual versus semi-automated assessment of breast density on standard mammograms. *Insights into Imaging*, 3(1):91–99, 2012.

[7] V. Murray, P. Rodriguez, and M.S. Pattichis. Multiscale am-fm demodulation and image reconstruction methods with improved accuracy. *IEEE Trans Im Proc*, 19(5):1138–1152, 2010.

[8] V. Murray, M. Pattichis, E. Barriga, and P. Soliz. Recent multiscale am-fm methods in emerging applications in medical imaging. *EURASIP J Adv in Signal Process*, 2012(1):23, 2012.

[9] A. Oliver, J. Freixenet, R. Marti, J. Pont, E. Perez, E. Denton, and R. Zwiggelaar. A novel breast tissue density classification methodology. *IEEE Trans Inf Technol Biomed*, 12:55 –65, 2008.

[10] S. Petroudi, T. Kadir, and M. Brady. Automatic classification of mammographic parenchymal patterns: A statistical approach. In *Proc of IEEE EMBC*, pages 798–801. IEEE, 2003.

[11] F. Schaefer, A. Waldmann, A. Katalinic, C. Wefelnberg, M. Heller, W. Jonat, and I. Schreer. Influence of additional breast ultrasound on cancer detection in a cohort study for quality assurance in breast diagnosis, analysis of 102,577 diagnostic procedures. *Eur Rad*, 20(5):1085–1092, 2010.

[12] J. Suckling, J. Parker, D.R. Dance, S. Astley, I. Hutt, C.R.M. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S.L. Kok, P. Taylor, D. Betal, and J. Savage. The mammographic image analysis society digital mammogram database. In *Second IWDM*, volume 1069, pages 375–378, 1994.

[13] J.N. Wolfe. Breast parenchymal patterns and their changes with age. *Rad*, 121:545–552, 1976.

# Developing A Simulation Toolbox For Biomedical Plenoptic Imaging

C Meah[1]
c.j.meah@pgr.bham.ac.uk
R Marshall[1]
r.j.marshall@pgr.bham.ac.uk
E Claridge[2]
e.claridge@cs.bham.ac.uk
K Bongs[3]
k.bongs@bham.ac.uk
I Styles[2]
i.b.styles@cs.bham.ac.uk

[1] PSIBS DTC
School of Chemistry
[2] School of Computer Science
[3] School of Physics and Astronomy
University of Birmingham
Edgbaston
UK, B15 2TT

**Abstract**

Plenoptic imaging allows a user to gain 3-D information from a single acquisition. A toolbox is presented with unified plenoptic simulation, rendering, and reconstruction, which allows the possibility for complete *in silico* exploration of potential designs and applications. The example of 3-D fluorescence microscopy is explored with plenoptic tomographic reconstruction.

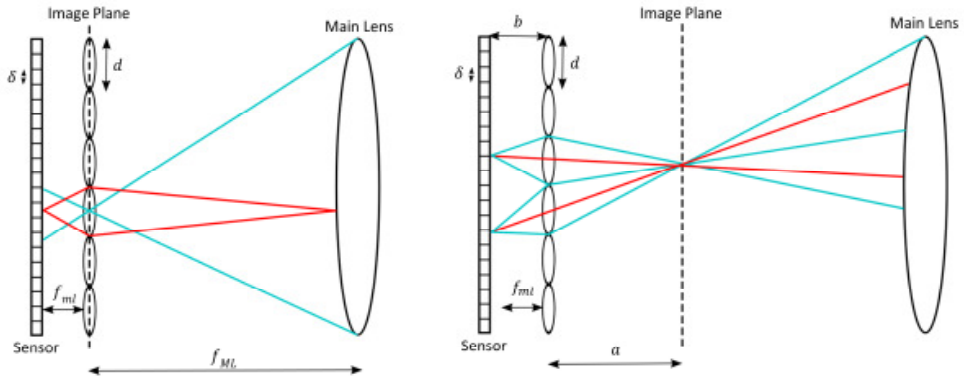
## 1 Introduction

Plenoptic, or light field, imaging is an exciting new technology which provides a user with 3-D information about a scene from a single camera in a single acquisition. Through post-processing of plenoptic data we can create a synthetic aperture, refocus digitally, generate novel viewpoints, and achieve 3-D reconstruction. These abilities have potential benefits in a wide variety of applications. In particular, medical imaging applications can benefit from this technology since it can dramatically reduce acquisition time. In translucent scenes, such as in microscopy, retinal imaging, or particle image velocimetry (PIV), this allows snapshot volumetric imaging which provides improved insight into dynamic micro-biology [6].

The main limitation of plenoptic imaging is the loss of spatial resolution, which is due to the spatio-angular trade-off [4]. Particular applications will require different resolutions, and so optimisation of plenoptic design is required for optimal imaging. This optimisation, however, is non-trivial and practical exploration can be expensive since many possible components must be available. Furthermore, placement of optical components must be accurate on the order of microns for optimal imaging to take place.

Simulation allows this exploration to occur without the need for a physical plenoptic system. This is not only beneficial for system design, but also for algorithm development

© 2014 . The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.




(a) (b)

Figure 1: A plenoptic camera utilises a microlens array to preserve the directional information of incoming rays. (a) shows the plenoptic 1.0 set-up, where the microlens array is placed on the main lens image plane. (b) shows the plenoptic 2.0 set-up, where the microlens array and sensor act as a relay system.

since validation and testing can be performed on well-characterised simulated data. This is important in biomedical imaging applications utilising 3-D segmentation and tracking.

This paper presents a unified, open source toolbox being developed to allow complete *in silico* plenoptic exploration. A camera can be designed, image acquisition simulated, and performance evaluated with in-built reconstruction algorithms. Virtual exploration is cheaper, easier, quicker, and more readily lends itself to quantitative evaluation. As an example of the tool's effectiveness, the software is utilised to investigate a plenoptic camera reconstruction of 3-D fluorescence microscopy.

# 2 Methods

A software tool has been written, using MATLAB (Release 2013a, The MathWorks Inc., Natick, MA, USA), which provides simulation and reconstruction methods for plenoptic data. This reconstruction can render images or 3-D visualisations from real or simulated plenoptic data. The simulation tool allows a user to design a bespoke virtual plenoptic camera, input a custom scene, and simulate the image acquisition to generate a raw image. The software is beneficial for optimisation of design, algorithm validation, and exploring applications.

## 2.1 Plenoptic Imaging

The light field is a term used to describe the distribution of light rays in space. A ray can be described by its position $(x,y,z)$ and direction $(\theta,\phi)$, giving a 5-D function often called the plenoptic function, $L_F(x,y,z,\theta,\phi)$. In free space, where there are no occluders, the radiance of a ray does not change along its length. This means we can lose one dimension, although this makes the parameterisation less obvious. We can characterise a ray by its intersection with two parallel planes (the $uv$- and $st$-planes), giving us a 4-D representation: $L(u,v,s,t)$.

A conventional camera only captures the positional information of light, but by placing a microlens array between the main lens and image sensor we create a plenoptic camera

and can capture the full 4-D function. The placement of the microlens array dictates the type of data captured, and therefore has bearing on the resolution achieved by the camera. Ng *et al*. developed a camera design with the microlens array on the main lens imaging plane, and the sensor one focal length behind the microlenses [9]. In this set-up, known as the *plenoptic 1.0* architecture, spatial resolution is determined by the number of microlenses, with the angular resolution dictated by the number of pixels under a microlens. This sacrifice in spatial resolution is a major drawback of plenoptic imaging. Lumsdaine and Georgiev traded back some of this angular resolution gain by adapting the camera design and creating a relay system with the microlens array and sensor [7]. In this set-up, known as the *plenoptic 2.0* architecture, the microlenses are a distance $a$ from the main lens image plane, and the sensor is a distance $b$ behind, where the distances obey the thin lens equation

$$1/f = 1/a + 1/b\,. \tag{1}$$

Now resolution is dictated by the $a/b$ relationship and is therefore more flexible. These two different cameras capture different data and therefore require different rendering algorithms in order to achieve optimal reconstruction.

In plenoptic 1.0 rendering, since each microlens captures information for one spatial point, we can simply sum the pixels underneath each microlens and tile them together. In terms of light field irradiance, the value at point $s,t$ can be calculated as

$$I(s,t) = \iint L(u,v,s,t)\mathrm{d}u\mathrm{d}v\,. \tag{2}$$

By applying $u$ and $v$ limits to the integrals, we can dictate the viewpoint and aperture size of the image produced. Refocusing, detailed in Ng *et al*. [9], can be achieved since rays can be traced from a virtual plane to the sensor, and we can interpolate the values that would have been captured at the refocusing plane.

In plenoptic 2.0 rendering, we can take multiple pixels from each microlens to contribute to a final rendered image rather than summing them together, meaning a dramatic increase in spatial resolution. The size of the window of pixels taken from each microlens dictates the focal depth at which objects are shown "in focus". Note "in-focus" in this case means an image void of plenoptic artifacts [3]. The effect of aperture can be incorporated into rendering by summing together information from neighbouring microlenses. Viewpoint is dependent on the difference between microlens and pixel window centres.

By utilising Equation 2 we can achieve these rendering effects from a single plenoptic data set. These are included functions in the toolbox presented, although a full description of the methods are beyond the scope of this paper (see [3][8] for details).

### 2.1.1 3-D Reconstruction

There are a number of methods available in order to gain a 3-D reconstruction from a plenoptic data set. A depth map representing the surface topography of a scene can be obtained by calculating the disparity between microimages [3].

Another method of 3-D reconstruction can be taken from the field of microscopy, where images focused on different image planes are acquired in order to produce a z-stack. We can attempt to eliminate contributions from out of focus planes in each image by performing deconvolution on this stack, yielding a 3-D model of the data [6]. This pipeline can be translated into plenoptic imaging since we can digitally refocus meaning a z-stack can be created post-acquisition.

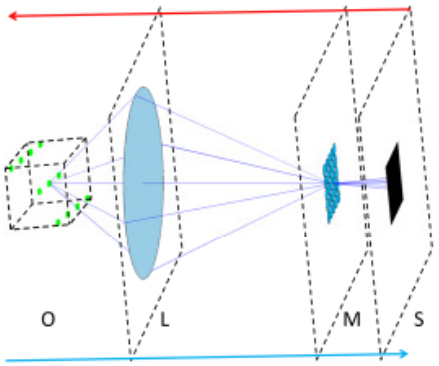



Figure 2: A 3-D schematic of the simulation. The input in this case is a volume representing fluorescent point sources in a transparent medium. Sample rays are shown traced through the system. The dashed lines represent (from left to right) the bounding box of the input volume (O), the main lens plane (L), the microlens array plane (M), and the sensor plane (S). The *red* arrow shows the direction of propagation for backwards simulations, and the *blue* for forwards simulation.

An alternative approach, which does not require the generation of a focal stack, is to pose the problem in terms of limited-angle tomography. Using techniques borrowed from computed tomography, we can iteratively solve the ill-posed and under-determined problem [1]. By treating measurements as projections through a volume, we can reconstruct voxel densities. This can be done separately for each "slice" of the volume by using the light field digital refocusing algorithm to project rays through slices [9]. The method outlined in [1] has been implemented in the reconstruction software so that a user can specify a volume of interest in object space. This is transformed into an image-side volume using the thin lens equation, seen in Equation 1, to calculate image distance ($d_i$) from object distance ($d_o$). The volume is also scaled with respect to magnification using $M = -d_i/d_o$. The light field can be projected through this image-side volume, and algebraic reconstruction techniques can be used to iteratively calculate the voxel densities.

## 2.2 Simulation

The simulation package allows a user the freedom to create a custom virtual camera. Ray transfer matrices (RTMs) [5] are used to propagate rays through an optical system. Ray transfer matrices are easily extensible, with the ability to account for intensity, and also incorporate more complex ray models [2]. Matrix notation allows the code to be highly vectorised which leads to improved software efficiency. Further speed improvements have been achieved through parallel computation using the MATLAB Parallel Toolbox$^{TM}$.

An optical component, such as a lens or microlens array, can be created by specifying parameters such as focal length, diameter, and thickness. For most simulations, the thin lens equation is sufficient for calculations, however a thick lens approximation is available if the radii of curvature are known for the lens. A component can be custom created so long as an operator matrix can be defined for it. A user can create an arbitrary system by placing optical components at specified positions.

We can combine optical component matrices to form a system matrix, meaning a single calculation can propagate a ray through a system. However, when considering a microlens

array we must first investigate which lens a ray hits to calculate how it is affected by that microlens.

The simulation can either propagate rays forwards or backwards through a system, with user specified ray space discretisation. Forwards simulation, the direction of which is shown in Figure 2 by the blue arrow, means that a ray originates from a point source and travels through the optical system where successful rays will hit a sensor pixel. This method depicts real-world imaging, but can be inefficient since there is no guarantee a ray will hit the sensor. The scene must be discretised into point source locations, with the level of discretisation affecting the quality of simulation. Backwards ray tracing, shown by the red arrow in Figure 2, means rays start at a sensor pixel and are traced until they hit an object in the scene. This is more efficient, particularly for dense scenes, but does not represent the physical process of imaging. Intersection points must be calculated for rays and objects in the scene which can be computationally intensive, however this can be alleviated through discretising space using an octree representation.

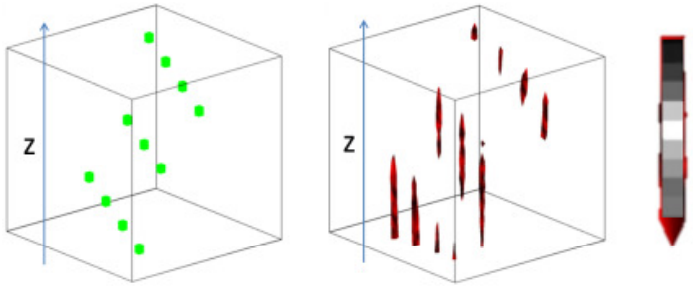
The input scene for the simulation software can be a voxel field or a triangulated mesh. The software allows for customising of meshes, for instance recolouring or combining meshes, which are treated like Lambertian surfaces. Meshes are widely used in computer graphics, and are good representations for many applications. The voxel field input is useful for representing isotropic point sources as found in confocal microscopy or PIV, although an option for user defined ray inputs allows for exploration of scenes which may result from more complex modelling such as monte-carlo simulations through scattering media.

To demonstrate the use of the software for investigating a plenoptic application we have simulated 3-D fluorescence microscopy. Point sources of equal intensity were virtually placed at varying positions in a volume representing optically transparent medium. A microlens array (pitch = 250$\mu$m, focal length = 500$\mu$m) was placed on the imaging plane of the main lens to which it was matched, and a 1 megapixel sensor was placed on the back focal plane of the microlens array.

# 3 Results and Discussion

The input volume and reconstruction for the application of 3-D fluorescence can be seen in Figure 3. For tomographic reconstruction the algorithm used to calculate voxel densities was based on Maximum Likelihood using Expectation Maximisation [10], although this choice of algorithm could be optimised for particular applications. It can be seen that the tomography performed on the plenoptic 1.0 data reconstructs the source locations in x-y plane accurately, but with an elongation in the z-direction. This can be attributed to the camera depth of field. Figure 3(c) shows an overlay of the voxel density values from the reconstruction. This shows that, with thresholding or further processing more definite object reconstructions could be estimated. The ability to accurately reconstruct tomographic data from a single plenoptic acquisition is an exciting prospect for practical experiments.

A new software toolbox for the simulation and reconstruction of plenoptic data has been created. The application of fluorescence microscopy has been demonstrated with tomographic reconstruction for point source localisation. The design of the plenoptic camera could be optimised by using the output from rendering methods as a metric, and optimisation of cameras and reconstruction algorithms for specific applications will be the basis of future work. More biomedical imaging problems could be addressed by incorporating specific light transport models into the input scene data. This means there is exciting poten-




(a) (b) (c)

Figure 3: Results from simulation. (a) The input scene of a volume containing 11 distributed point sources. The optical system for the simulation is above the volume (in the positive z direction shown by the arrow). (b) Isosurface of the tomographic reconstruction result. (c) An overlay of the reconstruction and the grayscale value showing the probability of density (white highest, black lowest).

tial to explore and optimise novel applications virtually which can inform future practical implementation.

# References

[1] T Fahringer and B Thurow. Tomographic reconstruction of a 3-d flow field using a plenoptic camera. In *42nd AIAA Fluid Dynamics Conference and Exhibit*, 2012.

[2] E Freniere et al. Edge diffraction in monte carlo ray tracing. In *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pages 151–157. International Society for Optics and Photonics, 1999.

[3] T Georgiev and A Lumsdaine. Reducing plenoptic camera artifacts. In *Computer Graphics Forum*, volume 29, pages 1955–1968. Wiley Online Library, 2010.

[4] Tl Georgiev et al. Spatio-angular resolution tradeoffs in integral photography. *Rendering Techniques*, 2006:263–272, 2006.

[5] A Gerrard and J Burch. *Introduction to matrix methods in optics*. Courier Dover Publications, 2012.

[6] M Levoy et al. Light field microscopy. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 924–934. ACM, 2006.

[7] A Lumsdaine and T Georgiev. The focused plenoptic camera. In *Computational Photography (ICCP), 2009 IEEE International Conference on*, pages 1–8. IEEE, 2009.

[8] R Ng. *Digital light field photography*. PhD thesis, stanford university, 2006.

[9] R Ng et al. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2(11), 2005.

[10] Y Vardi, LA Shepp, and L Kaufman. A statistical model for positron emission tomography. *Journal of the American Statistical Association*, 80(389):8–20, 1985.

# Integrated tools for next generation Bio-Imaging

Joaquin Correa[1,2]
JoaquinCorrea@lbl.gov
Manfred Auer[1,3]
MAuer@lbl.gov
David Skinner[1,2]
DESkinner@lbl.gov

[1] Lawrence Berkeley National Laboratory, Berkeley, USA
[2] National Energy Research Scientific Computing Center (NERSC), USA
[3] Life Sciences Division

**Abstract**

Exponential data scaling in Bio-Imaging, combined with the data fusion demands of increasingly multi-modal imaging approaches, require R&D in computational Bio-Imaging solutions that extend the ways in which scientists can leverage image based data. Bio-Imaging is leaving an era of images-as-results and entering an era where results come in the form of biological models built on big images and vast image collections. These models require scalable analysis, integration, and dissemination of image data delivered from advanced Bio-Imaging instrumentation.

In Big Data Bio-Imaging, images are the feedstock rather than the end result of the instruments we use to observe biological systems.

This work presents a web-based HPC-enabled one-stop-shop solution for producers and consumers of models built on imaging data.

# 1 Introduction

A fundamental problem currently for the biological community is to adapt computational solutions known broadly in data-centric science toward the specific challenges of data scaling in Bio-Imaging. This is of particular significance and urgency as recent developments in image data acquisition have moved the data set size from gigabytes per imaging session to terabytes, making it impossible for biologists to continue their common practise to manually analyse the data sets. In this work we target software solutions fit for these tasks, which leverages success in large-scale data-centric science outside of Bio-Imaging.

# 2 Architecture

## 2.1 Science gateways

A science gateway is a web-based interface to access high performance computing (HPC) resources and storage systems. Gateways allow science teams to access data, perform shared computations and generally interact with HPC resources over the web. Common gateway goals are to improve ease of use in HPC so that more scientists can

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

benefit from creating collaborative workspaces around data and computing and make data accessible and useful to the broader scientific community.

Science gateways can use a REST-based web API such as NEWT [7] to access HPC centers, including authentication, file management, job submission and accounting interfaces. These interfaces allow users to perform tasks through the web [8].

## 2.2 Image processing software stack

This work illustrates the deployment of custom-tailored image-processing algorithms and successful integration of a suite of universally used processing software among the biology community into an OMERO-based [1] science gateway [8] using HPC resources from the National Energy Research Scientific Computing Center (NERSC) [11].

The shared, scalable web service presented in this paper provides a modular and flexible one-stop-shop solution for producers and consumers of models built on imaging data by refining pixel data into actionable knowledge resources. As a result, this work offers a platform that involved co-design, has multi-modal capabilities, and is big-data ready.

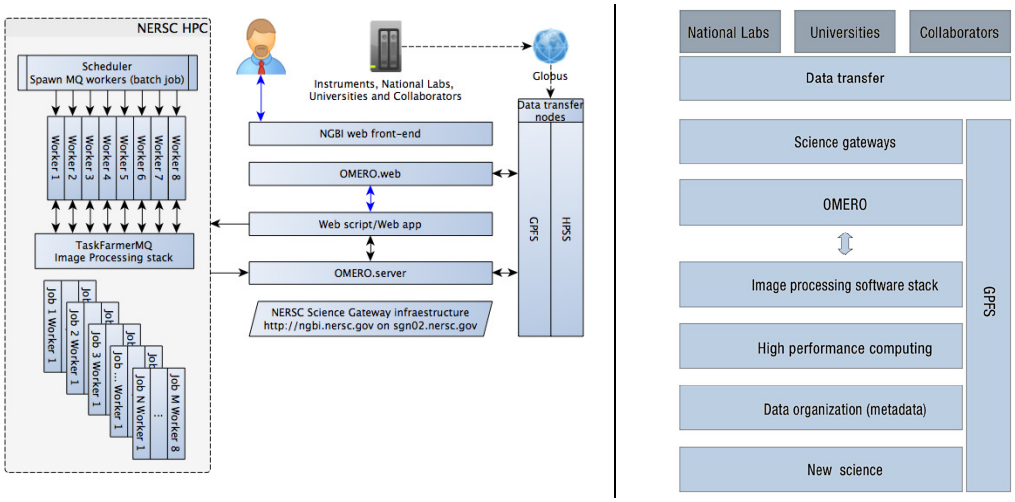
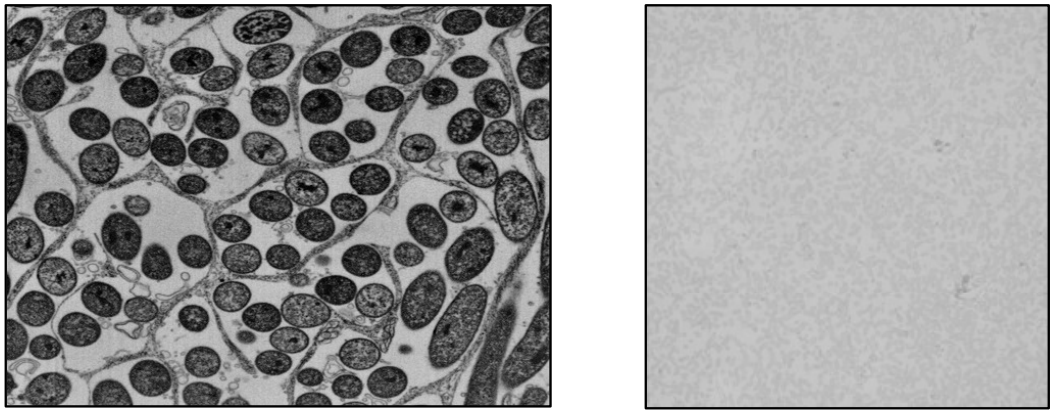



Figure 1: Architectural design. From top to bottom: (a) Data generation: National Laboratories, Universities and Collaborators (b) Data transfer to HPC facility (c) Web infrastructure: Science gateways [6] and OMERO [1] (d) Image processing stack: FIJI [3], Weka [3], Sci-kit learn [5], VLFeat [4], IMOD [9], CellProfiler [10] and UCSF Chimera [15] (e) Outcome: Data organization and New science.

# 3 Methodology

The authors demonstrate the capabilities of the platform presented in this work in the context of two monoculture bacterial biofilms, which are communities of bacterial cells with often-complex architectural organization. Studying biofilm organization is central to our understanding of microbial physiology [12], given that community organization represents the predominant bacterial lifestyle. Furthermore biofilms play an important

role for chronic infections and other microbial-mediated processes [13]. 3D datasets of Myxococcus xanthus [12] and Desulfovibrio vulgaris RCH1 (unpublished) microbial communities were used to illustrate segmentation, classification and visualization capabilities.



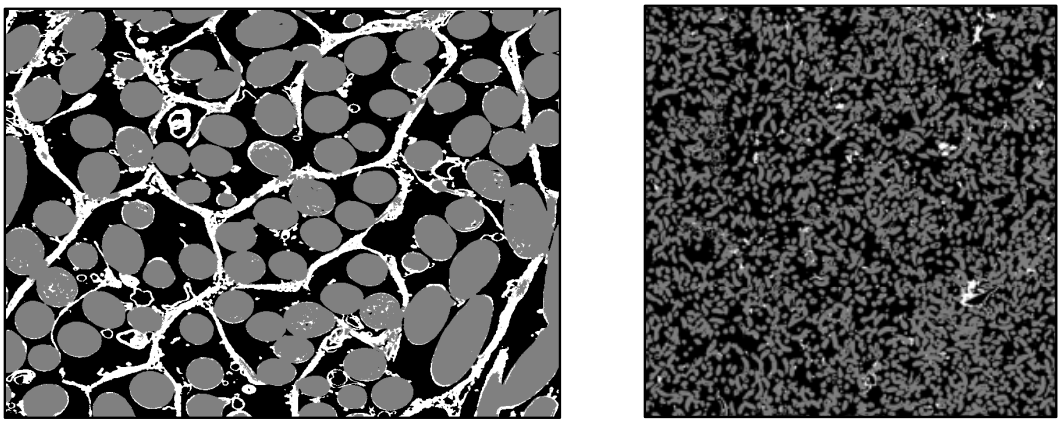
(a) (b)

Figure 2: Raw images: Showing a slice of the volume (a) Myxococcus xanthus (heavy metal stained) showing thin-walled partitions between cells (b) Desulfovibrio vulgaris RCH1 (unstained) showing extracellular metal deposits.

## 3.1 Segmentation and Classification

For the use cases the authors present; each dataset was segmented into three different classes (background, bacteria and biofilm), manually curated classifier models served as an input to the workflow.

Pre-processing enhancement operations were performed based on the image characteristics followed by image segmentation and classification achieved by using Weka [3] Fast Random Forest (FRF) implementation and a software layer that enables HPC by parallelizing serialized tasks.

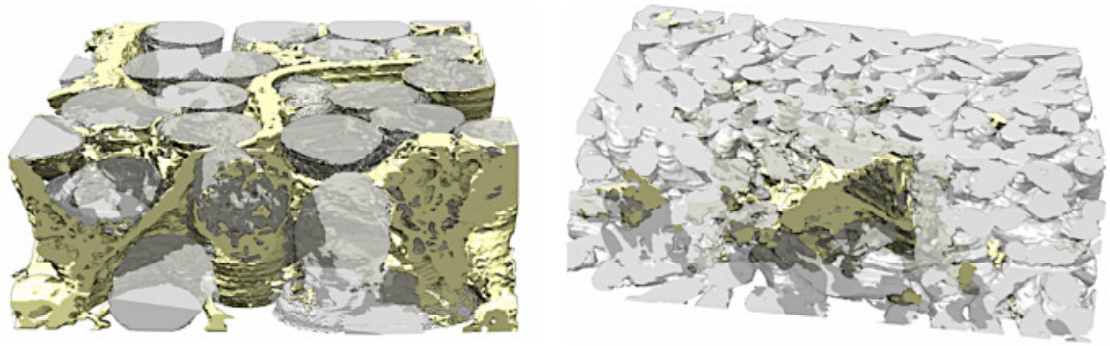


(a) (b)

Figure 3: Segmented datasets showing a slice of the volume: Background in black, bacteria in grey and partitions and metal deposits, respectively, in white (a) Myxococcus xanthus (b) Desulfovibrio vulgaris RCH1

### 3.2 Visualization

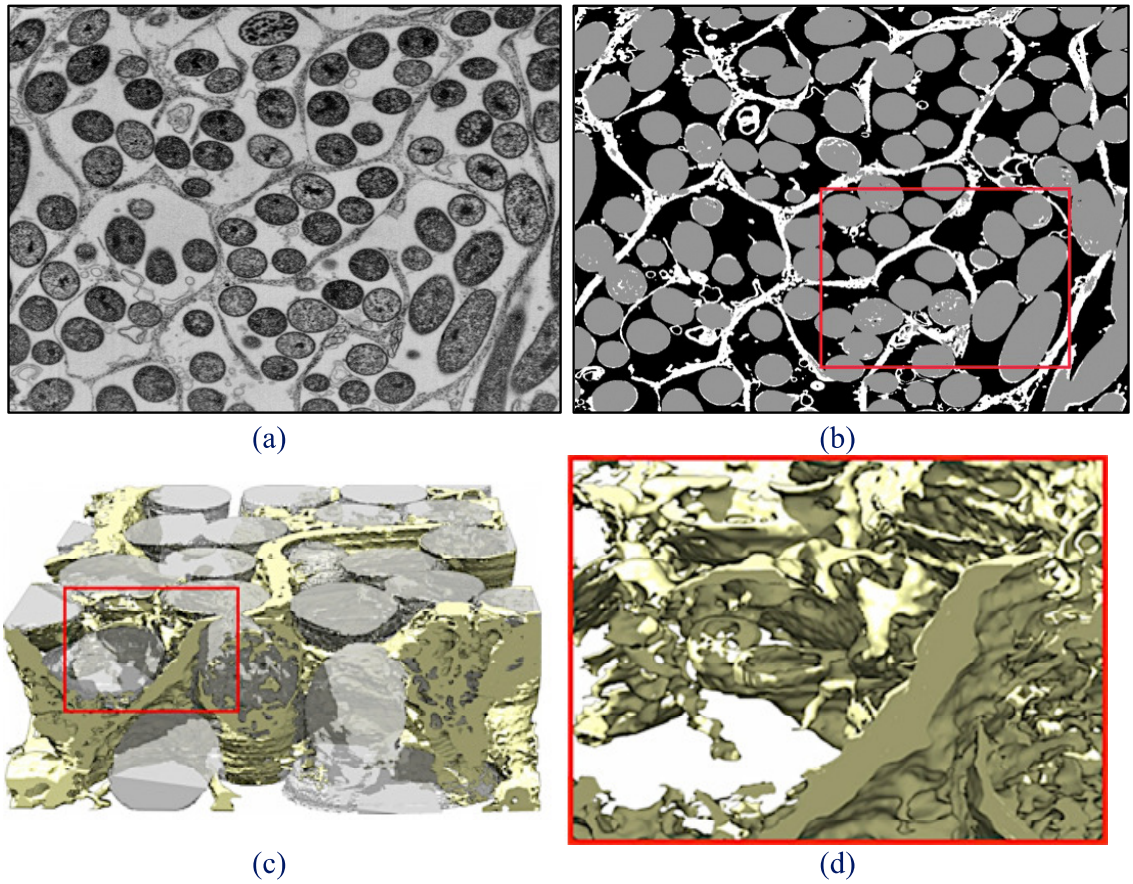
Corresponding segmented masks are used to produce 3D representation of the data using UCSF Chimera [15] as a running backend application accessed based on demand.



(a) (b)

Figure 4: Visualization of ROI with bacteria shown in grey and (a) partitions shown in yellow for Myxococcus xanthus or (b) extracellular metal deposits shown in yellow for Desulfovibrio vulgaris RCH1

## 4 Workflow summary

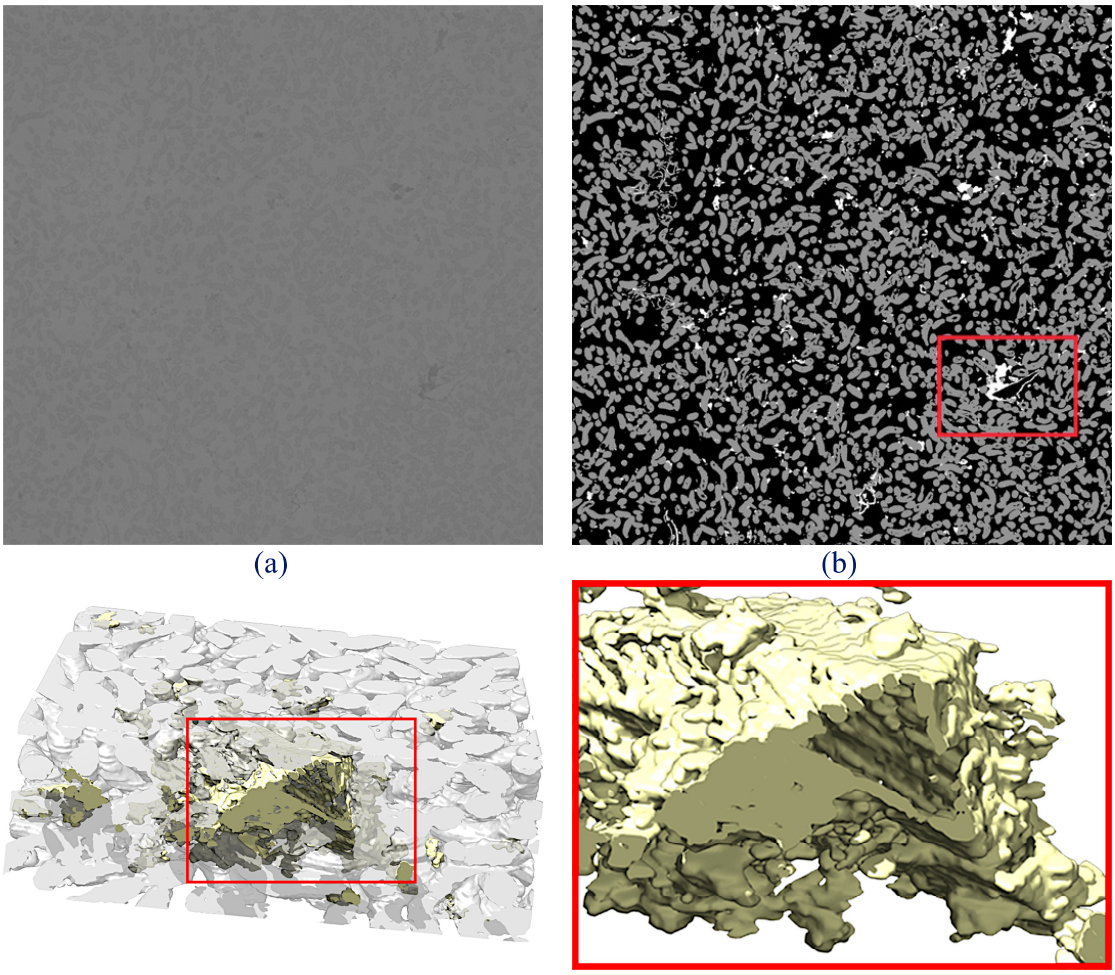


(a) (b)

(c) (d)

Figure 5: Myxococcus xanthus (a) Raw image showing a slice of the volume (b)

Segmented dataset showing a slice of the volume and ROI (c) 3D rendering of ROI showing bacteria and biofilm (d) higher magnification of ROI sub-region of (c) showing the 3D organization of the thin walls, with bacteria not being displayed.



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

Figure 6: Desulfovibrio vulgaris RCH1 (a) Raw image showing a slice of the volume (b) Segmented dataset showing a slice of the volume and ROI (c) 3D rendering of ROI showing bacteria and extracellular metal deposits (d) higher magnification of ROI sub-region of (c) showing the 3D organization of the metal deposits, with bacteria not being displayed

# 5 Conclusions

The platform we have presented provides new qualitative understanding of the 3D organization of the *Myxococcus xanthus* and *Desulfovibrio vulgaris RCH1* microbial communities [12] [13] and holds promise for image analysis of the big data flood, which is just beginning. Among the data challenges before the bio-imaging community are high-throughput, high-content and multi-modal image data, which are currently being produced at unprecedented scales.

# Acknowledgment

This work was supported by the Laboratory Directed Research and Development Program of Lawrence Berkeley National Laboratory under U.S. Department of Energy Contract No. DE-AC02-05CH11231.

# References

[1] G. J. Kleywegt, G. Zanetti, J. R. Swedlow, et al., OMERO: flexible, model-driven data management for experimental biology. *Nature Methods 9, 245–253,* 2012.

[2] K. Eliceiri, P. Tomancak, A. Cardona, et al., Fiji: an open-source platform for biological-image analysis, *Nature Methods 9(7): 676-682.* 2012.

[3] B. Pfahringer, P. Reutemann, I. H. Witten, et al., *The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.* 2009.

[4] A. Vedaldi and B. Fulkerson, VLFeat: An Open and Portable Library of Computer Vision Algorithms (2008), *URL: http://www.vlfeat.org/,* 2008.

[5] Pedregosa et al., Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830, 2011.

[6] NERSC, Bio-Imaging: High Performance Computing Facility Operational Assessment for the National Energy Research Scientific Computing Center 2013. 2014.

[7] NEWT, URL: https://newt.nersc.gov/

[8] National Energy Research Scientific Computing Center (NERSC) Science gateways, URL: http://www.nersc.gov/users/science-gateways/

[9] Kremer J.R., D.N. Mastronarde and J.R. McIntosh (1996) Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. 116:71-76.

[10] J. Moffat, P. Golland, DM. Sabatini, et al., *CellProfiler: image analysis software for identifying and quantifying cell phenotypes.* Genome Biology 7:R100. PMID: 17076895. 2006.

[11] National Energy Research Scientific Computing Center (NERSC), URL: http://www.nersc.gov/about/

[12] JP. Remis, JW. Costerton, M. Auer. Biofilms: structures that may facilitate cell-cell interactions. ISME J. 2010 Sep;4(9):1085-7.

[13] JD. Wall, LR. Krumholz (2006) Uranium reduction, Annu Rev Microbiol, 60: 149-166

[14] Integrated tools for Next Generation Bio-Imaging (NGBI), URL: http://ngbi.nersc.gov

[15] EF. Pettersen, TD. Goddard, CC. Huang, et al., UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004 Oct;25(13):1605-12.

Ahmad, Jahanzeb 203
Aichinger, Philipp 111
Aigner, Christoph 111
Al-Assam, Hisham 131, 143
Al-Baddai, Saad 223
Allinson, Nigel 99, 259
Al Okashi, Omar 143
Al-Subari, Karema 223
Argyriou, Vasileios 209
Auer, Manfred 283
Ayache, Nicholas 125

Bakas, Spyridon 73
Ballerini, Lucia 7
Barrington, Sally 173
Bischof, Horst 111
Bongs, Kai 277
Bourne, Tom 131
Bowyer, Simon 235
Boyer, Edmond 13
Brain, Keith 265

Carare, Roxana 167
Carey, Frank 67
Chatzimichail, Katerina 73
Cheng, Da-Chuan 197
Chupin, Marie 39
Claridge, Ela 277
Coats, Maria 67
Coenen, Frans 253
Colliot, Olivier 39
Constantinou, Ioannis 271
Correa, Joaquin 283
Cooper, Paul 241
Cootes, Tim 105
Criminisi, Antonio 125
Cury, Claire 39

Deghghani, Hamid 265
Denton, Erika 79
Denzler, Joachim 117
Dixon, Martin 179
Doulgerakis, Matthaios 209
Du, Hongbo 131, 143

Emrith, Khemraj 19

Fabritz, Larissa 265
Farren, Jessica 131
Fiorini, Samuele 7
Fisher, Mark 155
Flight, Rachel 241
Fordyce, Neil 59

Garcia-Finana, Marta 253
Glaunès, Joan 39
Grandas, Fabian Guiza 229
Griffin, Lewis 47
Groves, Chris 19
Gunn, Roger 4
Gunst, Jan 220

Hacker, Christian 53, 59
Hawkes, David 2
He, Wenda 79
Hecker, Dietmar 247
Hill, Derek 173
Hogg, Peter, 79
Holmes, Andrew 265
Huang, Tianjun 53
Hunter, Gordon 73

Jaccard, Nicolas 47
Jassim, Sabah 131
Janssens, Thomas 229
Juette, Arne 79

Kanthou, Chryso 235
Khadidos, Alaa 149
Khazendar, Shan 131
Kibiro, Minnie 79
Kim, Jaeil 33
Kirchhof, Paulus 265
Kolar, Radim 215
Körner, Marco 117
Krishna, Mahesh 117

Labbé, Bastien 73
Laissue, Philippe 179
Lambrou, Tryphon 99, 259
Landini, Gabriel 241
Lang, Elmar 223
Langlands, Kenny 143
Larsson, Jonas 139
Lašan, Michal 185
Lee, Daniel 125
Lelliott, Chris 143
Li, Chang-Tsun 149
Liu, Shing-Hong 197
Lohscheller, Joerg 247
Lucocq, John 53, 59

Maguire, Simon 143
Makris, Dimitrios 73
Margeta, Jan 125
Marsden, Paul 173
Marshall, Richard 277
McKenna, Stephen 53, 59, 67
Meah, Chris 277
Melvin, Dave 143
Milward, Michael 241

Namburete, Ana Ineyda 27
Nelson, Carl 179
Nixon, Mark 167
Nketia, Thomas 191
Noble, J. Alison 27, 191

Obara, Boguslaw 179
Odunaike, Solomon 91
Okuboyejo, Damilola 91
Olivo-Marin, Jean-Christophe 3
Olugbara, Oludayo 91
Ortmann, Wolfgang 117
Owen, Helen 229

Pansiot, Julien 13
Park, Jinah 33
Pattichis, Marios 271
Pattichis, Constantinos 271
Petroudi, Styliani 271
Phillips, Michael 173
Pierscionek, Barbara 209
Pizzorni-Ferrarese, Francesca 139
Podlipna, Petra 215
Poullis, Andy 19
Purwani, Sri 105

Qiu, Xinyi 33
Qureshi, Adnan 99, 161

Reveret, Lionel 13
Reyes-Aldasoro, Constantino 235
Rittsher, Jens 191
Rosado, Luis 85
Roesner, Imme 111
Ruggeri, Alfredo 7

Sahrim, Musab 167
Sanchez, Victor 149
Sayasneh, Ahmed 131
Schenk, Fabian 111
Selway, Joanne 143
Schetinin, Vitaly 161
Schick, Bernhard 247
Schuster, Maria 247
Shelton, Richard 241
Sidhu, Paul 73
Skinner, David 283
Slabaugh, Greg 19
Sluming, Vanessa 253
Smith, Lyndon 203
Smith, Melvyn 19, 203
Soltaninejad, Mohammadreza 99, 259
Soukup, Jindřich 185
Šroubek, Filip 185
Stebbing, Richard 27
Styles, Iain 241, 277
Süße, Herbert 117
Sun, Jiuai 203
Syeda, Fahima 265
Szita, Nicolas 47

Tomé, Ana Maria 223
Trucco, Emanuele 7
Twining, Carole 105
Tziakouri, Chrysa 271
Tzimiropoulos, George 209

Udomchaiporn, Akadej 253
Unger, Jakob 247
Urschler, Martin 111

Valdes-Hernandez, Maria 33
Van den Berghe, Greet 229
Vanhees, Ine 229
Vasconcelos, Maria João 85
Volberg, Gregor 223

Wang, Wenjia 155
Wardlaw, Joanna 33
Whiteford, Ian 33

Yang, Guang 259
Ye, Xujiong 99, 259
Yu, Ting Yue 265

Zhang, Jianguo 67
Zhang, Jingjing 67
Zhang, Lei 155
Zwiggelaar, Reyer 79

ISBN 978-1-901725-51-3