Medical Image Understanding and Analysis 2004

Proceedings of the eighth Annual Conference
Imperial College London, 23-24 September 2004

Edited by:

Daniel Rueckert
Jo Hajnal
Guang-Zhong Yang

Organisation:

Sponsored by: British Machine Vision Association (BMVA), Institute of Physics and Engineering in Medicine (IPEM), and the British Institute of Radiology (BIR).

Daniel Rueckert
Department of Computing
Imperial College London
South Kensington Campus
180 Queen's Gate
London SW7 2AZ, UK

Jo Hajnal
Imaging Sciences Department
Imperial College London
Hammersmith Hospital Campus
Du Cane Road
London W12 0NN, UK

Guang-Zhong Yang
Department of Computing
Imperial College London
South Kensington Campus
180 Queen's Gate
London SW7 2AZ

ISBN 1 901725 27 8

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 any means, with prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of the licenses issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

© BMVA 2004

The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant 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 in this book and cannot accept any legal responsibility for any error or omission that may be made.

Printed and bound in the United Kingdom at Imperial College London.

Medical Image Understanding and Analysis 2004

This is the eighth in a series of annual scientific and technical meetings designed to provide a UK forum for discussion and dissemination of research in medical image understanding and analysis. This year's meeting is organised by the Imaging Sciences Centre at Imperial College London and the Inter-disciplinary Research Consortium (IRC) "*From Medical Images and Signals to Clinical Information*" (MIAS). As in previous years, the meeting is supported by the British Machine Vision Association (BMVA), the Institute of Physics and Engineering in Medicine (IPEM), and the British Institute of Radiology (BIR).

The range and level of submission for this year's meeting has been of high quality. In contrast to previous years, authors were asked to submit a one-page abstract for review by the programme committee. We received 91 submissions, and all submitted abstracts were reviewed by three members of the programme committee. Based on their reviews and recommendations, 34 submissions were accepted for oral presentation and 35 submissions were accepted for poster presentation. The authors of papers which were selected for oral or poster presentation were then asked to submit a four-page paper which is included in these proceedings. We were particularly pleased to see the large number of high quality papers submitted by PhD students. At this year's MIUA, approximately 50% of all presentations are from PhD students.

Many thanks to all who helped in organising the event. In particular we would like to thank Michelle Ballmer, Barbara Claxton, Ann Halford and the members of the Visual Information Processing Group for their help in organising the conference. We would also like to thank David Risley for his excellent support of the conference management software, CAWS, which was used to manage the paper submission and delegate registration process. Finally, we are very grateful to the Programme Committee and, in particular, to all those who supported MIUA 2004 by submitting papers and attending the meeting. We hope that you enjoy the conference.

Daniel Rueckert
Jo Hajnal
Guang-Zhong Yang

Medical Image Understanding and Analysis 2004

Organisers and Co-chairs

Daniel Rueckert, Department of Computing, Imperial College London.
Jo Hajnal, Image Sciences Department, Hammersmith Hospital, Imperial College London.
Guang-Zhong Yang, Department of Computing, Imperial College London.

Steering Committee

E Berry, University of Leeds, IPEM
JM Brady, University of Oxford, RAE
JS Fleming, University of Southampton, BIR
DJ Hawkes, Kings College London
A Noble, University of Oxford
CJ Taylor, University of Manchester, BMVA

Programme Committee

SR Arridge, University College London
D Alexander, University College London
F Bello, Imperial College London
A Bhalerao, University of Warwick
E Claridge, University of Birmingham
T Cootes, University of Manchester
B Crum, King's College London
E Edwards, Imperial College London
D Hill, King's College London
A Houston, Royal Hospital Haslar
M Mirmehdi, University of Bristol
J Schnabel, King's College London
N Thacker, University of Manchester
A Todd-Pokropek, University College London
P Undrill, University of Aberdeen
R Zwiggelaar, University of East Anglia

Table of Contents

Session: Modelling

Modelling tissue deformations using free-form modes of vibration 1
J. A. Schnabel, P. G. Batchelor, C. Tanner, M. Quist, F. A. Gerritsen, D. J. Hawkes

A Matrix Framework for MR Motion Correction 5
P. G. Batchelor, D. Atkinson, D. L. G. Hill, J. V. Hajnal, D. Larkman

Building a 4D atlas of the cardiac anatomy and motion using MR imaging 9
D. Perperidis, M. Lorenzo-Valdés, R. Chandrashekara, A. Rao, R. Modiaddin, G.I. Sanchez-Ortiz, D. Rueckert

Decoupling of Respiratory Motion with Wavelet and Principal Component Analysis 13
F. Deligianni, A. Chung, G. Z. Yang

Session: Image Acquisition

Image-based Reduction of Artefacts in Radial Imaging 17
K. J. Lee, M. N. Paley, P. D. Griffiths, J. M. Wild

Noise Filtering and Testing for MR Using a Multi-Dimensional Partial Volume Model . . . 21
M. Pokrić, N. A. Thacker, D. C. Williamson

Consequences for imaging of the broadband nature of terahertz pulses 25
E. Berry, G. C. Walker, S. W. Smye, S. Reed, J. W. Bowen, R. E. Miles, A. G. Davies

Elasticity Reconstruction from Displacement and Confidence Measures of Ultrasound Image Sequence 29
J. Li, J. A. Noble

Session: Cardiac Image Analysis

Correction of misaligned slices in multi-slice MR cardiac examinations by using slice-to-volume registration 33
A. G. Chandler, T. Netsch, J. A. Schnabel, R. Razavi, D. L. G. Hill, D. J. Hawkes

A Coupled Active Contour Model for Myocardial Tracking in Contrast Echocardiography 37
N. Malpica, M. Ledesma-Carbayo, A. Santos, E. Pérez, M. A. García-Fernandez, M. Desco

Detection of Perfusion Defects using a Spatio-Temporal Analysis of Contrast Ultrasound Image Sequences 41
Q. R. Williams, J. A. Noble

Unsupervised Classification of Late Gadolinium Enhancement Cardiovascular MR 45
M. Ellington, X. P. Hu, A. John, D. Pennell, G. Z. Yang

Session: Segmentation

A Partial Volume 3D Gradient Magnitude Model 49
J. Chiverton, K. Wells, D. Johnson

Adaptive Contour Segmentation for Closely Packed Cytometric Assay 53
W. F. Clocksin

Automatic Segmentation of Liver from Computerised Tomography (CT) Images 57
A. Evans, T. Lambrou, A. Linney, A. Todd-Pokropek

Segmentation of Volumetric Prostate MRI Data Using Hybrid 2D+3D Shape Modeling . . 61
Y. Zhu, S. Williams, R. Zwiggelaar

Automatic segmentation of T1 parametric maps of breast MR images via a hidden Markov random field 65
G. Ketsetzis, M. Brady

Session: Registration

2 Dimensional Electrophoresis Gel Registration Using Point Matching and Local Image-Based Refinement 69
M. Rogers, J. Graham, R. P. Tonge

Whole Brain Voxel-based Analysis Using Registration and Multivariate Statistics 73
C. E. Thomaz, J. P. Boardman, D. L. G. Hill, J. V. Hajnal, A. D. Edwards, M. A. Rutherford, D. F. Gillies, D. Rueckert

Computing Covariances for Mutual Information Coregistration 77
P. A. Bromiley, M. Pokric, N. A. Thacker

Groupwise Non-Rigid Registration of Medical Images: The Minimum Description Length Approach 81
C. J. Twining, S. Marsland, C. Taylor

Session: Shape and Classification

Probabilistic shape analysis 85
J. G. Kim, J. A. Noble

A Generative Statistical Model of Mammographic Appearance 89
C. J. Rose, C. J. Taylor

Wavelet-Features for improved Tumour Detection in DCE-MRI 93
B. Lessmann, T. Twellmann, A. Degenhard, T. W. Nattkemper, P. Kessar, L. Pointon, M. Khazen, M. O. Leach

Visualisation of Breast Tumour DCE-MRI Data using LLE 97
C. Varini, T. W. Nattkemper, A. Degenhard, A. Wismueller

Automatic Capillary Measurement and Classification 101
K. Feng, P. D. Allen, T. Moore, A. L. Herrick, C. J. Taylor

Session: Neurological Image Analysis

Detailed Comparison of Non-Rigid Registration in Inter-subject Brain Registration 105
W. Crum, D. Rueckert, M. Jenkinson, D. Kennedy, S. M. Smith

Invisible differences in brain images of schizophrenics and controls 109
V. A. Kovalev, M. Petrou

Simulation of Realistic Abnormal SPECT Brain Perfusion Images: Methods and Applications 113
T. Ward, J. S. Fleming, S. M. A. Hoffman, P. M. Kemp

Three-Dimensional Voxel Segmentation of MR Brain Images with Deformable Model, Relative Fuzzy Classification and Spatial Affinity 117
R. Parveen, C. Ruff, D. McDonald, T. Lambrou, A. Todd-Pokropek

Session: Towards Clinical Applications

The effect of follicle volume measurement on clinical decisions 121
M. J. Gooding, D. Barber, S. Kennedy, J. A. Noble

Image-based Patlak analysis of dynamic 18FDG PET studies: incorporating recovery loss correction and blood curve normalisation to improve quantitative accuracy 125
H. A. Williams, P. J. Julyan, M. Ranson, J. Zweit, D. L. Hastings

Automated assessment of retinal image field of view 129
A. Fleming, S. Philip, K. Goatman, J. Olson, P. Sharp

Developing a Method of Automatic Prostate MRI Segmentation 133
P. D. Allen, J. Graham, D. C. Williamson, C. E. Hutchinson

Sessions: Posters

Detection and Characterisation of the Optic Disk in Glaucoma and Diabetic Retinopathy . 137
R. A. Abdel-Grafar, T. Morris, T. Ritchings, I. Wood

KCCA for fMRI Analysis 141
D. R. Hardoon, J. Shawe-Taylor, O. Friman

Comparison of methods for visualising changes from orthognathic surgery 145
L. Miller, E. Berry, D. O. Morris

Combined Ultrasound Speckle Pattern Similarity Measures 149
J. Revell, M. Mirmehdi, D. McNally

Comparison of Segmentation Methods for Cytometric Assay 153
C. Restif, W. F. Clocksin

Spatio-Temporal Free-Form Registration of Cardiac MR Image Sequences 157
D. Perperidis, R. Mohiaddin, D. Rueckert

Stereo-photogrammetry and image analysis for Fetal Alcohol Syndrome screening 161
T. S. Douglas, E. M. Meintjes, C. L. Vaughan, R. Grobbelaar, D. L. Viljoen

Automating cell analysis: intracellular particle tracking 165
A. Fitch, F. Suess, P. Courtney

Image-based identification of hibernating myocardium 169
N. M. I. Noble, D. L. G. Hill, M. Breeuwer, R. Razavi

Creation of patient-specific CFD models by morphing a previously-meshed reference geometry using image registration 173
D. M. Jones, D. R. Hose, P. V. Lawford, D. L. G. Hill, R. S. Razavi, D. C. Barber

Modelling Normality and Functionality of the Lungs 177
H. Singh, M. Crawford, J. Curtin, R. Zwiggelaar

An image retrieval approach for thermal medical images 181
G. Shaefer, S. Y. Zhu, B. Jones

MAP model of simultaneous segmentation and registration 184
C. Xiaohua, M. Brady, D. Rueckert

2 Dimensional Electrophoresis Gel Registration Using Point Matching and Local Image-Based Refinement 188
J. Ivins, A. Turpin

Correcting differential intensity inhomogeneity 192
E. B. Lewis, N. C. Fox

Combining Topological and Geometric Features of Mammograms to Detect Masses 196
B. W. Hong, M. Mellor, S. Soatto, M. Brady

Tracking of myocardial walls and study of contractility and thickening 200
R. Casero-Cañas, J. Timperley, J. A. Noble, H. Becher

Towards Classification of Prostate MRI 204
R. Zwiggelaar, Y. Zhu, S. Williams

Partitioning the Cingulate Gyrus using Bezier curves 208
R. G. Boyes, J. Barnes, N. C. Fox

Registering 3D Lung Surfaces Using the Shape Context Approach 212
M. Urschler, H. Bischof

Eccentric Elliptical Contours in Total Hip Replacements 216
S. Kerrigan, S. J. McKenna, I. W. Ricketts, C. Wigderowitz

Feasibility study of a novel techniques for constructing respiratory motion models, for use in 4D lung cancer radiotherapy planning 220
J. R. McClelland, A. G. Chandler, J. M. Blackall, S. Ahmad, D. B. Landau, D. J. Hawkes

A comparison of serial MR neuroimaging at 1.5T and 3T 224
B. Sneller, J. Vickers, M. A. Dresner, R. A. Heckemann, J. Fitzpatrick, M. E. Miquel, J. V. Hajnal, D. L. G. Hill

A MR neuroimaging tissue segmentation study at 1.5T and 3T 228
J. Vickers, B. Sneller, J. V. Hajnal, D. L. G. Hill, R. Heckeman, S. M. Smith

Improved T1 estimation in contrast enhanced MRI for pathology detection using optimum flip angle acquisitions 232
G. Ketsetzis, M. Brady

Generation of multi-phase (4D) CT data based on a nonrigid voxel intensity registration approach and MR breath hold scans 236
Y. Bouchareb, R. Janka, W. A. Kalender

3D MR liver modelling for use in pre-surgical planning 240
H. Ibrahim, M. Petrou, K. Wells, S. Doran, O. Olsen

Dynamic MRI with regular under-sampling for increased time resolution: x-f choice 244
S. Malik, S. Schmitz, J. V. Hajnal

An Improved ASM Approach Using Mixture Models to Prostate Segmentation from MR Images 248
Y. Zhu, S. Williams, R. Zwiggelaar

Label propagation in MRI: potential use in diagnostic imaging 252
R. A. Heckemann, J. V. Hajnal, R. I. Aviv, C. J. Harvey, D. L. G. Hill, D. Rueckert

Predicting the effect of image blur on boundary shape 255
E. Y. T. Ho, C. J. Taylor

Orientation Coherence Optimisation in Tensor Image Registration 259
K. M. Curran, D. C. Alexander

Cellular Algorithms for Medical Image Processing 263
L. Ibáñez

Flaws in Existing Spatio-Temporal and Temporo-Spatial Realignment of FMRI Data 267
P. R. Bannister, J. M. Brady, M. Jenkinson

Volumetric analysis of brain MRI using Geometry Deformable Models 271
W. W. Lee, I. Richardson

Modelling tissue deformations using free-form modes of vibration

Julia A Schnabel[a*], Philipp G Batchelor[a], Christine Tanner[a],
Marcel Quist[b], Frans A Gerritsen[b], David J Hawkes[a]

[a]Imaging Sciences Division, Guy's, King's and St. Thomas' Medical School, King's College London, UK
[b]Research and Advanced Development, Medical IT, Philips Medical Systems Nederland BV, NL

Abstract. In this work we propose to compute modes of vibration from finite element methods (FEM) for free-form deformations (FFDs) in order to model biomechanical tissue motion and deformations.

1 Introduction

Since physical modes of vibration were first introduced to computer vision applications for correspondence matching [1], their potential has been explored for various medical image analysis tasks, including motion tracking, surgical navigation and shape classification. In this work, we investigate whether modes of vibration can provide a low-dimensional and compact representation of biomechanically plausible types of deformation for their use in in non-rigid image registration.

2 Method

Modes of vibration are computed by decoupling the FEM system equilibrium equation, $\mathbf{R} = \mathbf{M}\ddot{\mathbf{U}} + \mathbf{K}\mathbf{U}$, into a basis defined by the $\mathbf{M}$-orthonormalized eigenvectors of $\mathbf{M}\mathbf{K}^{-1}$. Here, $\mathbf{R}$ is the load vector, $\mathbf{U}$ the global displacement vector, and $\mathbf{K}$ and $\mathbf{M}$ the global stiffness and mass matrix. This yields a set of eigenvectors Φ and eigenvalues Λ, which describe the image-embedded object's generalized nonlinear axes of symmetry $\mathbf{K}\Phi = \mathbf{M}\Phi\Lambda^2$. For constant masses, the generalized eigenproblem reduces to a standard one.

For free-form modes of vibration, the global system and mass matrices can be assembled from the local elements defined by a regular mesh of free-form deformations (FFDs) using B-splines [2]. Material properties such as Young's modulus E and Poisson's ratio ν can be associated with each FFD element e, yielding local element stiffness matrices $\mathbf{K}^{(e)}$ using Gauss quadrature integration, which are assembled to the global system:

$$\mathbf{K} = \sum_e \mathbf{K}^{(e)} = \int_{V^{(e)}} \mathbf{B}^{(e)T}\mathbf{C}^{(e)}\mathbf{B}^{(e)} dV^{(e)} \tag{1}$$

where $\mathbf{C}$ is the strain material matrix, and $\mathbf{B}$ is the strain displacement transformation matrix. The FFD model can then be reparameterized into a frequency-ordered description in form of a linear combination of the modes:

$$\mathbf{T}_{local} = \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3} B_l(u)B_m(v)B_n(w)(\mathbf{b}\Phi)_{i+l,j+m,k+n} \tag{2}$$

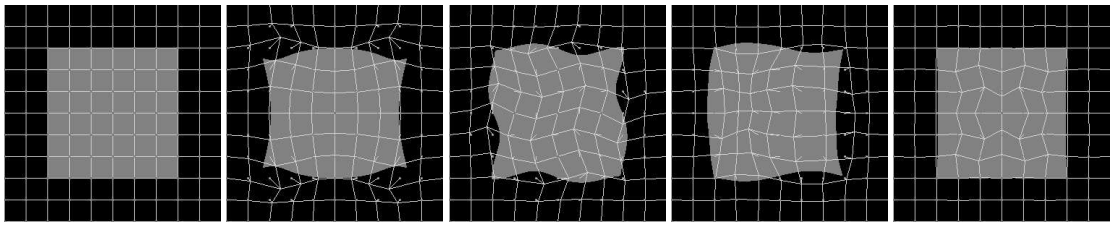
where $\mathbf{b}$ is a parameter vector weighting the basis of eigenvectors Φ, thus describing how each control point moves within each mode. All modes corresponding to rigid-body modes, background elements, or fixed degrees of freedom (DOFs) have zero eigenvalues, and can be excluded from the eigen solution *a priori* using efficient sparse matrix methods. Additionally, less relevant high frequency modes can also be excluded. Hence, our proposed modal reparameterization of FFDs can reduce the number of DOFs, while constraining the model to physically motivated and plausible types of deformations.

3 Results

We have computed free-form modes of vibration on a set of Magnetic Resonance (MR) image volumes, including MR mammography, MR liver volumes, and the brainweb1 phantom from the Montreal Neurological Institute, as well as synthetic images. Figure 1 shows a synthetic example image with superimposed free-form deformation modes at increasing frequencies of vibration, illustrating the effect of gross-to-fine deformation modelling.

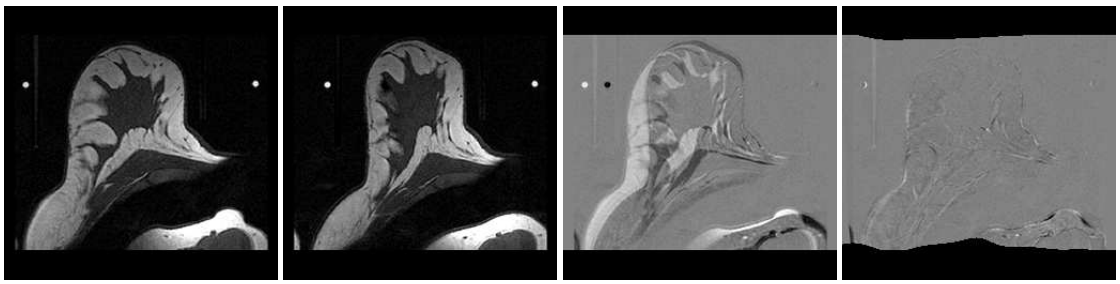
In the following, the exemplar MR mammography reconstruction will be described. Figure 2 shows example slices for MR mammography data of a volunteer undergoing a controlled compression in a breast biopsy coil [3], as well as subtractions before and after non-rigid registration using the algorithm described in [2].

*Corresponding author: julia.schnabel@kcl.ac.uk



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

Figure 1. Synthetic example image containing one square object, with a Young's modulus of $E = 100Pa$. (a) Relaxed state. (b)–(e) Example modes at increasing frequencies of vibration.



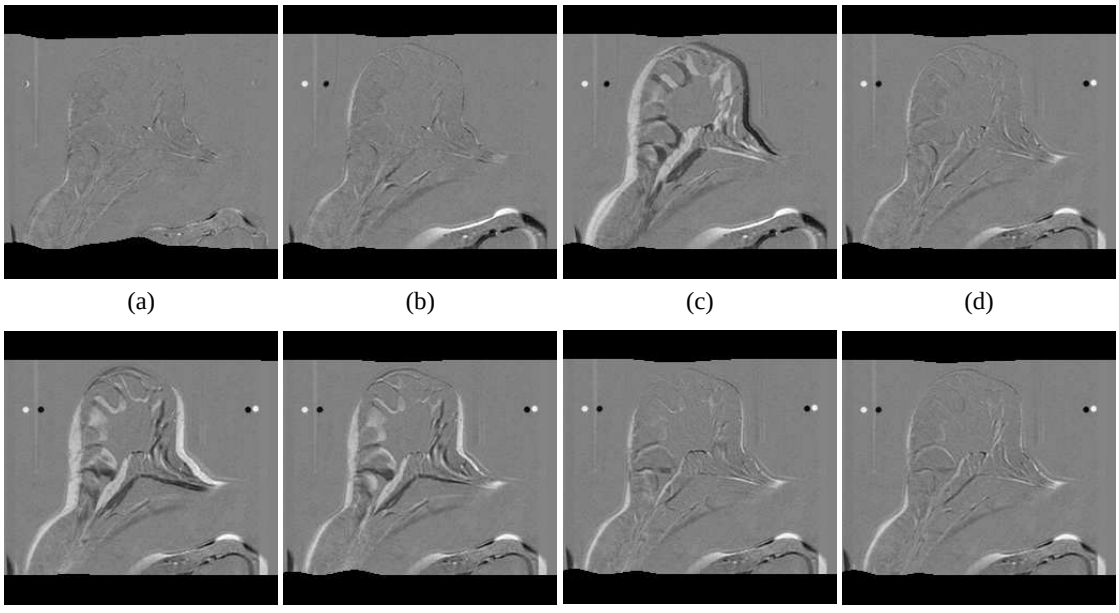
(a) (b) (c) (d)

Figure 2. Example axial slice through MR mammography volumes of a volunteer undergoing a controlled compression in a breast biopsy coil [3]. (a) Before compression. (b) After compression. (c) Subtraction (a)-(b). (d) Subtraction after volumetric non-rigid registration.

Figure 3 shows the corresponding subtraction slices after modal reconstruction of the non-rigid deformation field, using all available modes or only a mode subset, taking either all displacements or only displacements within the breast tissue into account. The subsets of modes were selected from those modes with strictly positive frequencies of vibration. While a reconstruction with all modes, using all displacements or only breast tissue displacements (Fig. 3 (a) and (b)), captures all or a large proportion of the deformations (with an RMS error of 6×10^{-15} and $0.55mm$, respectively), it is desirable to truncate the modal basis further to contain only the modes of non-zero frequency of vibration. In this example, these modes amount to just over a quarter (1098 of 4131 modes) of the total number of modes. However, as Fig. 3 (c) illustrates, using only the "positive" modes for reconstruction leaves a large residual rigid motion, which needs to be corrected for, e.g. using a closed form solution (see Fig. 3 (d)). To further reduce the positive modal basis, only the first N positive modes can be used (Figs. 3 (e)–(h)), yielding a decreasing reconstruction error for increasing N. In fact, for this example, using less than 10% of the total number of modes (or just over a third of the positive modes) (Fig. 3 (g)) yields submillimetre reconstruction accuracy.

For a further numerical analysis of the reconstruction error, Fig. 4 plots the RMS reconstruction error curves over all modes, as well as the scaled frequencies of vibration λ. Zero frequencies contribute to either rigid body motion or background motion, which is reflected by a constant reconstruction error using only foreground deformations (reconstructions R2, with an initial fall-off in Fig. 4 (a) for the rigid body modes). Using only positive frequency modes allows to reconstruct only the foreground displacements, hence the lower reconstruction error when ignoring background deformation (R4) as opposed to taking all deformations into account (R3). Using additional rigid motion correction (Fig. 4 (b)) shows an almost equivalent error when using only positive modes (R4) compared to all modes (R2), both based on foreground displacements only. Reconstruction $R1$ over all deformations and modes finally shows the complexity of the registration task, with a slowly decreasing rather than constant error for zero-frequency modes. In a registration application, where a modal reconstruction due to lack of a gold standard deformation field cannot be performed, these modes, if taken out of the modal basis, will have to be compensated for by rigid body motion correction.

As two further demonstrations for free-form modes of vibration, and on-going work, Figs. 5 and 6 show vibration modes for a slice of the MNI brainweb1 phantom, and a reformatted slice of an MR liver volume of a volunteer at maximum exhale position. For the former, vibration modes from a high-resolution 2D FFD ($5mm \times 5mm$) with three tissue components (grey and white matter, and CSF) were computed, using Young's moduli of, $E = 4kPa; 8kPa; 0kPa$, respectively, yielding 3096 modes, of which the first non-rigid modes are shown. For the liver example, 3D modes were computed for a one tissue component model ($E = 1kPa$). We are currently investigating



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

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

Figure 3. Example subtraction slice of volunteer MR mammography data after volumetric modal reconstruction of the displacements obtained from non-rigid registration (see Fig. 2 (d)). Top row: using (a) all 4131 modes, (b) like (a) but only displacements within the breast tissue, (c) all "positive" modes (1098 modes) and only displacements within the breast tissue, and (d) like (c) but using additional rigid motion correction. Bottom row: like (d) but using only the first N positive modes for (e) $N = 10$, (f) $N = 81$, (g) $N = 385$, and (h) $N = 675$.

reconstructive properties by means of deformation simulations and gold standard registrations.

Discussion and Conclusions

In our initial experiments, reparameterising free-form deformations to a modal basis of vibrations has shown to have a big potential to incorporate biomechanical tissue properties into a transformation models in an elegant way, while providing means to constrain the searh space in a registration scenario to physically plausible types of deformation. A further prospect of this approach is to use only a releavnt subsect of non-rigid free-form modes of vibration in order to recover lower frequency (gross) deformations and to reduce the parameter search space even further. We are currently investigating the effect of varying material property values, multi-resolution FFDs and associated modes, as well as comparing the free-form modes of vibration against gold standard FEM models.

Acknowledgements

Julia Schnabel and Philipp Batchelor have received funding from Philips Medical Systems. Christine Tanner has received funding from EPSRC GR/M52779 and MIAS-IRC.

References

1. A. Pentland & S. Sclaroff. "Closed-form solutions for physically based shape modelling and recognition." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13(7)**, pp. 715–729, 1991.
2. D. Rueckert, L. I. Sonoda, C. Hayes et al. "Non-rigid registration using free-form deformations: application to breast MR images." *IEEE Transactions on Medical Imaging* **18(8)**, pp. 712–721, 1999.
3. C. Tanner, A. Degenhard, J. A. Schnabel et al. "A comparison of biomechanical breast models: a case study." In M. Sonka & J. M. Fitzpatrick (editors), *Proc. SPIE Medical Imaging 2002: Image Processing*, volume 4684, pp. 1807–1818. SPIE, 2002.

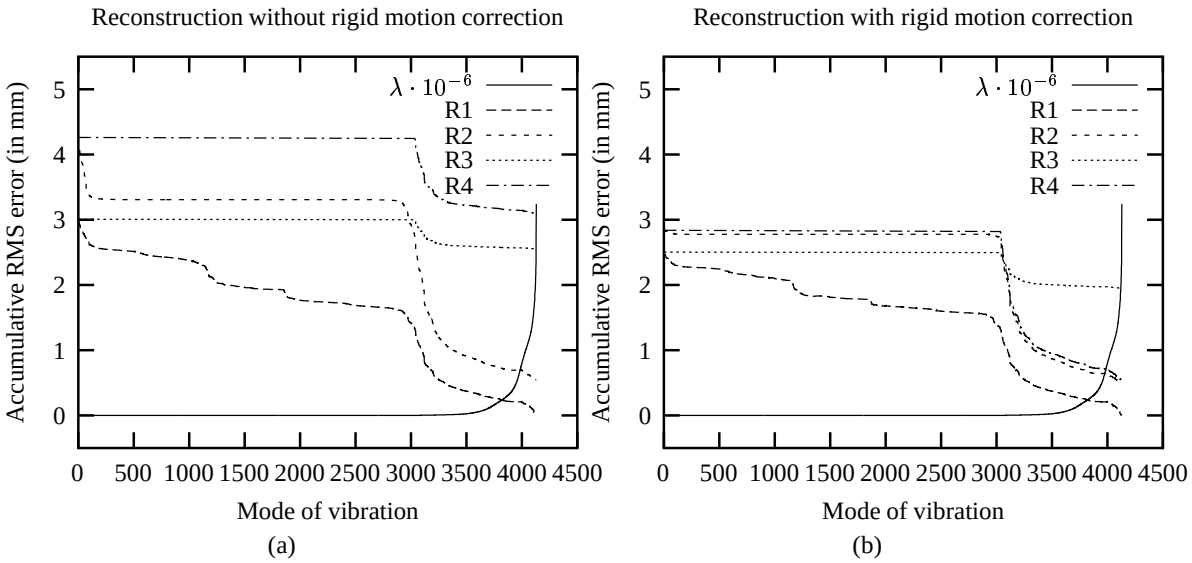


Figure 4. Accumulative modal reconstruction error of volumetric tissue displacements obtained from non-rigid registration, and corresponding frequency of vibration λ (scaled by 10^{-6} for better display purposes). (a) Error without additional rigid motion correction. (b) Error with additional rigid motion correction. R1: all modes; R2: all modes and only foreground displacements; R3: all modes with positive eigenvalues; R4: all modes with positive eigenvalues and only foreground displacements.

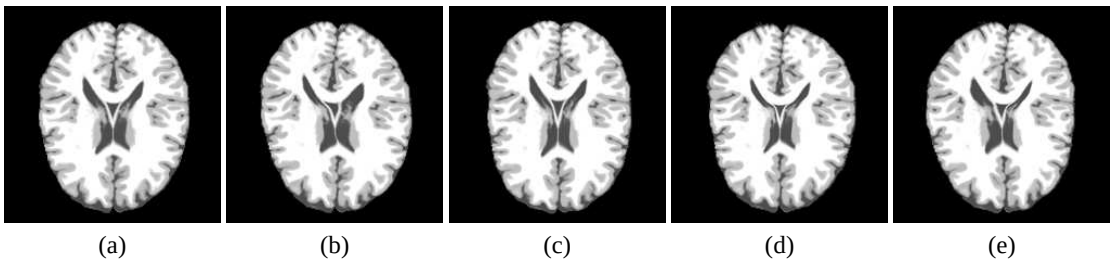


Figure 5. MNI brainweb1 slice and example modes of vibration. (a) Relaxed state. (b-e) First non-rigid modes.

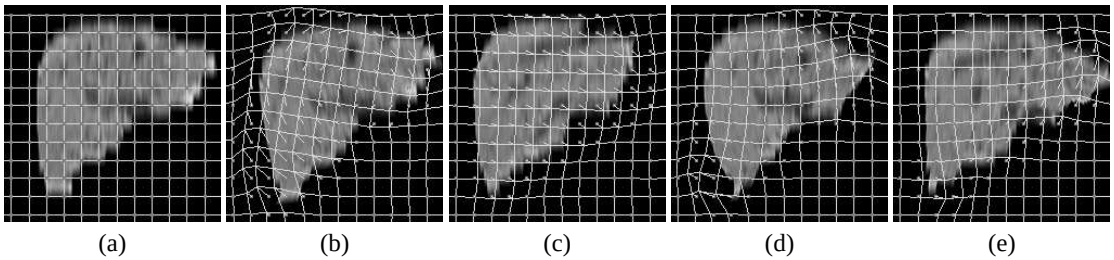


Figure 6. Reformatted sagittal slice through MR liver volunteer data at maximum exhale and example modes of vibration. (a) Relaxed state. (b)–(d) Rigid body and non-rigid modes.

A Matrix Framework for MR Motion Correction

P. G. Batchelor[a*], D. Atkinson[a], D. L. G. Hill[a], J. V. Hajnal[b], D. Larkman[b]

[a]Imaging Sciences Division, King's College London, 5th floor Thomas Guy House
[b]Imaging Sciences Department, Imperial College London

1 Introduction

We present a new framework to describe the effect of motion upon MR images. We establish a large matrix equation giving the corrupted image directly from the ideal object. We then show that pseudo-inversion of this system is possible by efficient linear algebraic methods when the motion is known. We concentrate here on the issue of the general motion correction when the motion is assumed known, but we also describe methods to find this motion.

2 Theory

Images acquired in the Fourier domain, and which have been motion corrupted are in general corrected by using two simple facts about Fourier transforms: the Fourier transform of a translated image is the original data with a phase shift, and the Fourier transform of a linearly transformed image is the Fourier transform of the original image, evaluated at the inverse transformed positions, magnified by a determinant. Two problems appear when trying to generalise to non-affine motions, or motion acquired in multiple shots. First, the Fourier data acquired at different times, i.e., k-space positions, is not consistent, which causes problems for rotations for example. Secondly, no general statement holds for the Fourier transform of an image corrupted by nonrigid motion, in terms of transforms of the coordinates in the underlying space: we can't simply find a nonrigid motion of k-space corresponding to a nonrigid motion of image space. Here, we show how to bypass these problems. The idea is to work with what is observed, namely the images or their Fourier transform, and not on the abstract underlying x- or k-spaces. We do this by expressing every operation as linear operations on images s interpreted as long vectors.

2.1 Known Motions

Spatial transformation matrices. Spatial transformations induce image transformations on the images whose domain (FOV) is transformed. In a discrete setting, the space of $N = n_1 \times n_2$ images $\mathrm{Img}(n_1, n_2)$ is a linear space of dimension $n_1 \cdot n_2$ in which images are arrays s. For an image s, and a transformation $\mathbf{x} \mapsto \mathbf{x} + \vec{u}(\mathbf{x})$ the transformed image satisfies $s_{\vec{u}}(\mathbf{x}) = s(\varphi_{\vec{u}}^{-1}(\mathbf{x}))$. The boundedness of the field of view makes these induced maps non-invertible in general. An image whose *signal* stays within the field of view, however, can be reconstructed.

Fourier, subsampling, and aliasing matrices. Fourier transformations are very fast thanks to Fast Fourier Transform algorithms (FFT), which perform the transformation without building explicitly a dense matrix. Nevertheless, for theoretical discussions, and for our purpose, it is useful to introduce the matrix of the 2D Fourier transform F_{n_1,n_2}. It is a unitary matrix, whose inverse is its hermitian transpose $F^H_{n_1,n_2}$. Subsampling of a vector can also be represented by a matrix, with ones in the diagonal at the corresponding positions, zero otherwise. We call the matrix corresponding to subsampling at time t A_t. Aliasing is the consequence of subsampling in the image domain, thus $a_t := F^H_{n_1,n_2} A_t F_{n_1,n_2}$ is the matrix which maps the image data to the aliased image data. For rectangular field of views, and undersampling along one spatial dimension (Phase Encode PE), the computation factors along dimensions, and the matrices have a block structure. Suppose the aliasing takes every n_s ($N_s := N/n_s$) line, with an offset at the start of t, $t = 0, \ldots, n_s - 1$ (i.e., line t, $t + n_s$, $t + 2n_s$, ...).

$$(a_t)_{mn} = \begin{cases} \frac{1}{n_s} e^{2\pi i t \frac{m-n}{n_1}} & \text{when } (m-n) \text{ is a multiple of } N_s \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

All the aliasing matrices have the same regular sparsity pattern, the magnitude of the nonzero components are just $1/n_s$, but the nondiagonal parts (which cause the ghosts) have phases which when summed over all shots lead to cancellations.

*Correspondence: {philipp.batchelor,derek.hill}@kcl.ac.uk

Motion in k-space. Suppose u_t describes the spatial transformation at time t, and A_t the sampling of k-space at this time. The observed image s is related to the ideal object image s_0 by the sequence: —Spatially transform s_0 to S ($s_0 \mapsto u_t s_0$); —Fourier transform ($s_0 \mapsto F_{n_1,n_2} u_t^- s$); —Extract the lines corresponding to shot t ($s_0 \mapsto A_t F_{n_1,n_2} u_t^- s$); —Set these lines in S ($s_0 \mapsto S := \sum_t A_{t'} F_{n_1,n_2} u_{t'}^- s$); —Let s be the inverse Fourier transform of S ($s \mapsto s := \sum_t F_{n_1,n_2}^H A_t F_{n_1,n_2} u_t^- s$). We have obtained a simple matrix form for the motion corruption from the expressions in brackets: the motion corrupted image is a superposition of aliased views of the object in different positions,

$$s = \left(\sum_{t=0}^{n_s-1} a_t u_t \right) s_0 =: \boldsymbol{\gamma} s_0. \tag{2}$$

We call this matrix the *ghosting* matrix (hence the choice of notation, we use the Greek form to avoid confusion with the g-factor from parallel MR). Implicitly, when the motions are known, inversion of this matrix should recover an object's image from its motion corrupted view, *whatever* the motion and time-sampling pattern in k-space. It is, however, unlikely that the matrix will be invertible, for the reasons given above. Note that it is also large, of size $n_1 n_2 \times n_1 n_2$, *but*, the matrix-vector multiplication by $\boldsymbol{\gamma}$ can be implemented very efficiently with FFT and subsampling. Nevertheless, its importance should not be underestimated, as it gives a benchmark to compare other faster, but approximate methods. Note that we have chosen the image space version of the equation, as it is images that are visualised. A k-space version of equation (2) would be $S = (\sum_t A_t U_t)\, S_0 =: \boldsymbol{\Gamma} S_0$ where $U_t := F_{n_1,n_2} u_t F_{n_1,n_2}^H$ is the k-space representation of motion. One important remark is that we are able to express the effect of any motion on the Fourier transform of an image, thus represent motion in k-space, but expressed on the *observed* images S, not on the coordinates.

Approximate inverses. Let us consider the following algorithm , on an image s, supposed to result from motions u_t occurring during the sampling. This is an exact version of the first method mentioned in the Introduction: Spatially transform s by $\varphi_{t'}^{-1}$ into s_t, ($s \mapsto u_{t'}^- s$); —Fourier transform s_t to $S_{t'}$, ($s \mapsto F_{n_1,n_2} u_{t'}^- s$); —extract the lines corresponding to shot t ($s \mapsto A_{t'} F_{n_1,n_2} u_{t'}^- s$), and set these lines in S ($s \mapsto S_1 := \sum_t A_{t'} F_{n_1,n_2} u_{t'}^- s$). Let s_1 be the inverse Fourier transform of S_1, an approximation to s_0. Then $s \mapsto s_1 =: \boldsymbol{\gamma}^- s := \sum_t F_{n_1,n_2}^H A_{t'} F_{n_1,n_2} u_{t'}^- s$. This algorithm is 'wrong' in the sense that in general s_1 will be different from s_0. The reason is that the shot lines corresponding to t were indeed transformed by a known transform, but only these lines, and this algorithm applies the transformation to the whole image. Ideally, the correct transformation would be applied only to the correct lines. This, however, would always assume that the other lines have been corrected, thus to correct the lines from time t, one needs the lines *corrected* at every other time t', but to get these, one would need the lines corrected at t, etc. In other words, the problem of the algorithm is that it is sequential, whereas the exact solution can only be achieved simultaneously, by matrix inversion. Spatial transformations whose transform matrices commute with aliasing matrices are then special in that they allow for a fast exact reconstruction, using for example this algorithm. This is true for translations along aliased lines. Even for more general motions, if the displacements are small, the matrices $a_{t'}$ and u_t^- might approximately commute, in the sense that $a_{t'} u_t^- u_t^- a_{t'}$ has small norm.

Inverses. Algorithms such as the one above are empirical. The advantage of knowing the explicit matrix $\boldsymbol{\gamma}$ is that the search for inversion can be made in a much more systematic way. Of course, the matrix being singular, this statement should be relativised, but, as we already mentioned, the non-regularity is caused by the loss of portions of the image out of the field of view. In other words, we know what the kernel of the matrix is, and it suggests that the singularity might not be too much of a problem, as the images we are interested in have small intersection with this kernel. Hence, instead of heuristic guesses of the inverse, mathematical constructions of *quasi-inverses* might be possible. Here we use the term quasi-inverse of a matrix A to mean a matrix A^- such that the error $\|z - A^- A z\|$ is small, in a sense defined by the problem, in particular for 'reasonable' z. The conjugate gradient method is the basis of many modern methods for the solution of large linear systems, and a derived Krylov subspace algorithm can be constructed for least-squares systems. For a regular matrix the conjugate gradient algorithm would always converge to the unique solution (in a finite number of steps). One feature is that it uses only matrix vector product operations, and we have a very efficient implementation of this using FFT and image transforms. Thus, such an algorithm works even for a singular matrix, in the sense that it will converge to *some* solution. (In a different context of medical imaging, this property of the conjugate gradient for singular matrices was used in [1]).

2.2 Unknown Motions

We have now described what to do if we know the motions. Of course the remaining question is how to find this motion. This question has already been studied in some detail, see for example ([2–4]), but quite often the methods are limited to specific types of motion, bulk translations, sometimes rotations ([4]), which depend on few parameters. The other extremes are some transformations used in nonrigid registrations which sometimes depend on thousands of parameters. A template for motion correction will be

1. Generate a guessed motion.
2. Transform the image by this guessed motion.
3. Assess this image, and improve the guess (return to 1.)

3 Results

We first show the difference between the methods in Fig. 1, of the motion corrupted a), to the empirical in b), and the image corrected by matrix inversion (here by explicitly building the sparse matrix). Clearly, the empirical method does not give a correct reconstruction of the image (here corrupted in two shots, by a relatively large rotation). We also illustrate the method on a 3D dataset, showing that dimension is not a huge limitation, here for

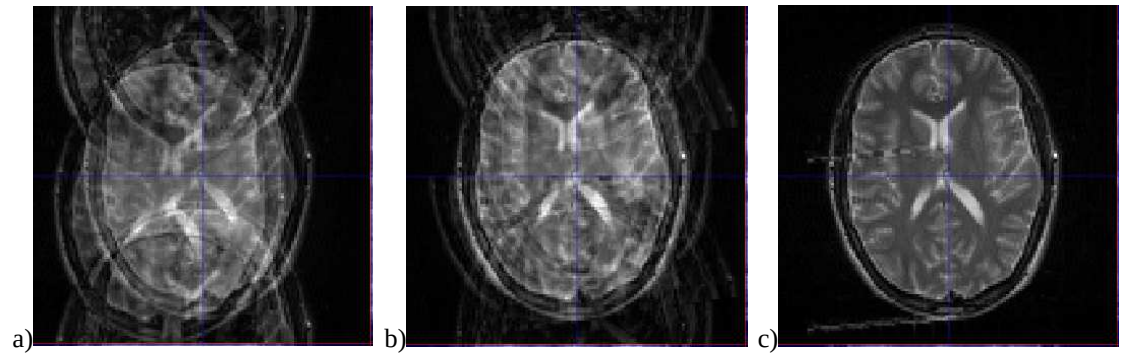


Figure 1. Assessment of approximate algorithm for rotations: a) Original image; b) the reconstruction by the approximate inverse, showing clearly that for large rotations, this approximate method does not give good results; c) the solution of the linear system, by conjugate gradient.

random affine motions in four shots, and even a nonrigid motion (Fig. 2). Finally, in Fig. 3 we show an example

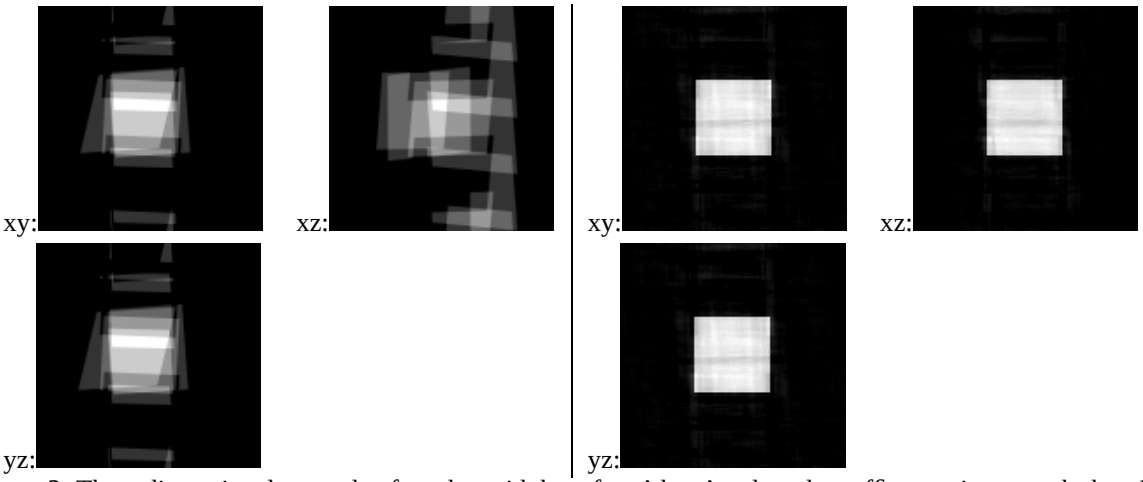


Figure 2. Three dimensional example of a cube, with here four 'shots' and random affine motion at each shot. We use here a simple cube as it allows to see more clearly the effects. Three panel views, left motion corrupted, right corrected.

of a nonrigid motion, here the radial deformation $\rho \mapsto \rho_0(\rho/\rho_0)^{\alpha_t}$ where ρ is the distance from the centre of the image, r_0 a fixed radius, and α_t controls the deformation, here $\alpha_0 = \alpha_3 = 1$ and $\alpha_1 = \alpha_2 = 1.5$.

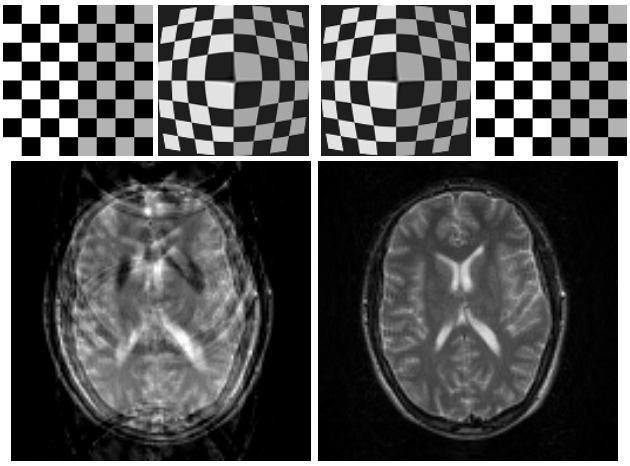


Figure 3. Nonrigid example, which simulates a pulsation where the image deforms according to the views in the first row, to simulate a pulsation. Left: motion corrupted, right: corrected.

4 Conclusion

We have established a formula for motion in MR which does not restrict the type of motion, or the sampling pattern. Motion corrected images obtained using derived optimisation strategies show very promising results. The approximate inverse corresponds to the traditional correction in k-space, but the difference between Figs. 1-b and 1-c, shows clearly that for large motions the correct matrix inversion is much better. In summary, correction of very general type of motion becomes a realistic prospect. The motion can be known, from navigators, or markers for example. The method we described applies directly in this case, otherwise an optimisation strategy is required, where the matrix framework is applied at each iteration. This framework is an advance on previous work in the complexity of the motion it can correct, and the computational efficiency of the algorithm.

In future work we will consider the whole motion correction problems, in particular the problem of finding unknown motions. Note that the correct motion correction for known motion will clearly produce better guesses than approximate methods, thus is likely to improve the convergence (see our earlier work in [2]).

Acknowledgements

We thank the EPSRC for grants no GR/N04867 and AF/001381, and the EPSRC/MRC funded ”Medical imaging and Signals” IRC. We also thank Pablo Irarrazaval and Klaas Pruessman.

References

1. S. Angenent, S. Haker, A. Tannenbaum, and R. Kikinis. On the Laplace-Beltrami operator and brain surface flattening. *IEEE Trans. Med. Imag.*, 18(8):700–711, 1999.
2. D. Atkinson, D. J. Larkman, J. Hajnal, P. G. Batchelor, and D. L. G. Hill. Coil-based artifact reduction. *Magn. Reson. in Med.*, to appear, 2004.
3. K. P. McGee, A. Manduca, J. P. Felmlee, S. J. Riederer, and R. L. Ehman. Image metric-based correction (autocorrection) of motion effects: Analysis of image metrics. *J. Magn. Reson. Imag.*, 11:174–181, 2000.
4. J. G. Pipe. Motion correction with PROPELLER MRI: Application to head motion and free-breathing cardiac imaging. *Magn. Reson. in Med.*, 42:963–969, 1999.

Building a 4D atlas of the cardiac anatomy and motion using MR imaging

D Perperidis[a]*, M. Lorenzo-Valdés[a], R. Chandrashekara[a],
A.Rao[a], R. Mohiaddin[b], G.I. Sanchez-Ortiz[a], D. Rueckert[a]

[a]Visual Information Processing Group, Department of Computing, Imperial College, London,
[b] Cardiovascular MR Unit, Royal Brompton Hospital, Imperial College, London

Abstract. In this paper we describe the construction of 4D atlas of human heart using cardiac MR imaging. This probabilistic atlas captures the cardiac anatomy and function of a healthy heart. In order to build the atlas we have acquired tagged as well as untagged MR image sequences from 11 healthy volunteers. The untagged MR image sequences for each subject are segmented and then mapped into a common reference coordinate system using a novel spatio-temporal registration algorithm to produce a 4D probabilistic model of the cardiac anatomy. In addition, the tagged MR image sequences are used to derive motion fields between the end-diastolic and the end-systolic frames which describe myocardial contraction patterns in each subject. These motion fields are also mapped into our spatio-temporal reference coordinate system to produce a 4D statistical model of cardiac function.

1 Introduction

Cardiovascular diseases are the single most important cause of death in the developed world [1]. Their early diagnosis and treatment is crucial in order to reduce mortality and to improve patients' quality of life. Magnetic resonance imaging (MR) is playing an increasingly important role for the high resolution imaging of the cardiovascular system. The area of computational anatomy is an active research area. Most of the work in in this area focuses on atlas based approaches. The majority of medical atlases focus primarily on the anatomy and function of the human brain. Early attempts to construct atlases of the human brain have focused of models built from a single subject. However, these atlases do not contain any information regarding the anatomical and functional variability across the entire population. In order to address this problem, probabilistic atlases have been developed which include information from a set of subjects making them more representative of the normal population [2]. Recently, Mazziotta et al. presented a 4D atlas and reference system of the brain which includes macroscopic and microscopic information about the structure and function of the human brain in a large population across different age ranges [3].

In recent years a large number of approaches have been developed for the volumetric-based modeling of the heart. An excellent review of these approaches can be found in Frangi et al. [4]. While there have been a number of attempts to build statistical shape models of the cardiac anatomy [5, 6] very little attempts have been made to build a computerized atlas which captures both the anatomical and functional variability of the heart across a group of subjects.In this paper, we present a atlas of the cardiac anatomy and motion using MR imaging. We have developed anatomical and functional probabilistic atlases for the left ventricle of the heart, the myocardium and right ventricle. Furthermore, we have also developed a greylevel atlas of the cardiovascular system. The following chapters describe the approach for building the atlas and provide illustrations of the atlas.

2 Constructing the cardiac atlas

In order to build the atlas the following steps are used:

- The untagged cardiac image sequences are segmented into left and right ventricle as well as myocardium using an automated segmentation algorithm.
- The untagged cardiac image sequences are mapped into a common spatial and temporal coordinate system using a 4D registration algorithm.
- The myocardial motion fields for each subject are calculated from the tagged MR images using a motion tracking algorithm.
- The myocardial motion fields are mapped into the common spatial and temporal coordinate system using the transformation calculated previously.
- A 4D probabilistic map for each anatomical structure is calculated. Furthermore, 4D statistical motion model is calculated.

*email:dp1@doc.ic.ac.uk

2.1 Segmentation of the cardiac image sequences

The method developed by Lorenzo-Valdés et al. [7] has been used to segment the image sequences. In this method the first frame of each image sequence is segmented manually and then the segmentation is propagated to the subsequent frames using a non rigid registration algorithm. The images sequences are segmented into four anatomical structures: the left ventricle, the myocardium, the right ventricle and the background.

2.2 Mapping the image sequences to the same spatio-temporal coordinate system

In order for the cardiac atlas to describe the average anatomy and function of the heart all image sequences have to be mapped into the same spatio-temporal coordinate system. A spatio-temporal registration method has been used in order to align the image sequences into the same reference system. This method is similar to one which we have previously introduced [8]. It uses a 4D mapping which has been resolved into decoupled spatial and temporal components $\mathbf{T}_{spatial}$ and $\mathbf{T}_{temporal}$ respectively where

$$\mathbf{T}_{spatial}(x,y,z) = (x'(x,y,z), y'(x,y,z), z'(x,y,z)), \mathbf{T}_{temporal}(t) = t'(t)$$

One consequence of this decoupling is that each temporal frame t in image sequence I will map to another temporal frame t' in image sequence I', ensuring causality and preventing that different spatial regions in a 3D image $I_k(x,y,z)$ will be warped differently in the temporal domain by $_{temporal}$.

The spatial transformation used is an affine transformation with 9 degrees of freedom which account for spatial differences caused by orientation, translation and scaling. To compensate for differences in the length of the cardiac cycle and for differences in the length of the contraction and relaxation phases across different subjects we have chosen a temporal transformation which consists of a global part and a local part:

$$\mathbf{T}_{temporal}(t') = \mathbf{T}^{global}_{temporal}(t) + \mathbf{T}^{local}_{temporal}(t)$$

$\mathbf{T}^{global}_{temporal}$ is an affine transformation which scales the image sequences in order to match the end-systolic and end-diastolic time points . $\mathbf{T}^{local}_{temporal}$ is modeled by a free-form deformation using a 1D B-spline. To define a spline-based temporal free-form deformation we denote the temporal domain of the image sequence as $\Omega_t = \{(t) \mid 0 \leq x < T\}$. Let Φ denote a set of n_t control points ϕ_t with a temporal spacing δ_t. Then, the temporal free form deformation can be defined as a 1D cubic B-spline:

$$\mathbf{T}^{local}_{temporal}(t) = \sum_{l=0}^{3} B_l(u)\phi_{t_{i+l}} \tag{1}$$

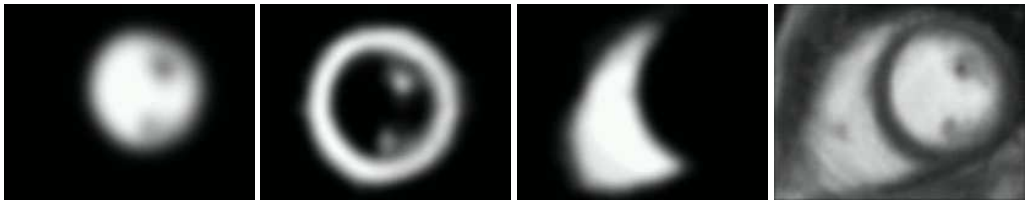
where $i = \lfloor \frac{t}{n_t} \rfloor - 1$, $u = \frac{t}{n_t} - \lfloor \frac{t}{n_t} \rfloor$ and B_l represents the l-th basis function of the B-spline. $\mathbf{T}^{local}_{temporal}$ deforms the temporal characteristics of each image sequence in order to follow the same motion pattern with the reference image sequence. The optimal spatial and temporal transformation is found by maximising a voxel based similarity measure, the normalised mutual information (NMI) [9]. The NMI of two image sequences can be calculated directly from the joint intensity histogram of the two sequences over their spatio-temporal domain of overlap. During the optimisation new voxel values are generated in the temporal domain using linear interpolation and trilinear interpolation in the spatial domain. The optimisation is carried out using an iterative downhill descent algorithm to calculate the optimal transformation.

2.3 Analysis of the myocardial motion

There are a number of different techniques to extract myocardial motion from tagged MR images. To construct the cardiac motion atlas the myocardial motion patterns in all subjects are analysed using the method developed by Chandrashekara et. al. [10]. In this method the deformation of the myocardium is calculated from the end-diastolic frame to the end-systolic frame and is obtained by non-rigid registration of each frame of the tagged MR image sequence to the end-diastolic frame. In order to map the motion fields into the common reference coordinate system we use an approach introduced by Rao et al. [11]. In this approach, the motion fields of different subjects are mapped into the same coordinate system using a vector field transformation technique which accounts for differences in the size, orientation and shape of the hearts. After mapping all the motion fields to common coordinate system the average motion as well as the principal modes of variation of the motion patterns can be calculated.

2.4 Building the probabilistic atlas

After calculating the transformations to map each image sequence to the common spatio-temporal coordinate system we can use these transformations to map the segmented image sequences into the same spatio-temporal



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

Figure 1. Examples of the 4D atlases of (a) probabilistic atlas of the left ventricle (b) the myocardium, (c) the right ventricle, and (d) the greylevel atlas. Animations of the atlases are found at `http://www.doc.ic.ac.uk/~dp1/Conferences/MIUA04/ATLAS/.`

coordinate system. The atlas of each anatomical structure is formed by averaging the segmented image sequences of the corresponding anatomical structure. Before producing the average image sequence, each segmented image sequence is blurred with a Gaussian kernel. The use of blurring during the construction of the atlas is needed due to the low out of plane resolution which results in significant partial volume effects in the segmentation. Blurring the images with a Gaussian kernel addresses this problem by modeling this uncertainty in the tissue classification. Furthermore, we are currently using a small population to build our atlas. This may lead to undesirably sharp transitions in the tissue probability maps. This effect will be significantly reduced by the addition of more subjects in the atlas.

3 Materials

In order to build the 4D atlas we have acquired tagged and untagged MR image sequences from 11 healthy volunteers. All untagged MR image sequences were acquired on a Siemens Sonata 1.5 T scanner using TrueFisp pulse sequence in form of a series of short-axis images covering the entire left ventricle from base to apex. The short-axis images had a resolution of 192×256 with a pixel size of $1.48mm \times 1.48mm$ and a slice thickness of 10mm which (which is the typical pixel size used in these acquisitions). Shape-based interpolation was used in order to resample the spatial domain of the segmented image sequences and produce more detailed representations with isotropic voxels of 1mm^3 size. For all subjects, the tagged MR image sequences were acquired in the same scanning session as the untagged images. For the tagging, short- and long-axis images were acquired to produce true 3D motion fields. The tagged short-axis images have been acquired with the same geometry as the untagged short-axis images. A cine breath-hold sequence with SPAMM tag pattern was used.

4 Results and discussion

Figure 1 provides examples of the probabilistic atlases in form of tissue probability maps and average greylevel intensity image. The smaller the intensity value the less the probability of the particular pixel to belong to the particular structure (the intensity 255 corresponds to probability of 1, while the intensity 0 corresponds to probability of 0). In figure 1 (a) and (b) we can clearly see that the papillary muscles of the myocardium are more blurred than the rest of the image. This is due to the fact that there is large variability in the position and the size of the papillary muscles. The registration cannot correct differences in the shape of the structure which results to high degree of blurring.

We have also produced volume renderings of the combined tissue probability and motion maps. The volume rendering was produced using the *Visualisation Toolkit* package (www.vtk.org). Figure 2 shows the volume rendering of the atlases. In figure 2 (a),(d) we can see the left ventricle with the average deformation of the myocardium from two different views, while in (b),(e) we can see the myocardium atlas from the same two different views and, finally, in (c),(f) we can see the volume rendering of the atlas of the right ventricle. The lower the probability of a voxel to belong to a structure the more transparent the voxel is rendered.

5 Conclusions

We have presented a 4D atlas of the heart using cardiac MR images. The atlas is created from cardiac image sequences from 11 healthy volunteers. We have built probabilistic atlases for the left ventricle, the myocardium and the right ventricle and also a grey-level atlas of the entire heart. These atlases represent the anatomy and the function of a healthy heart . For future work we are planning to include more subjects in the atlas. The more subjects the atlas has the more accurate representation of the population is. We also planning to use different spatio-

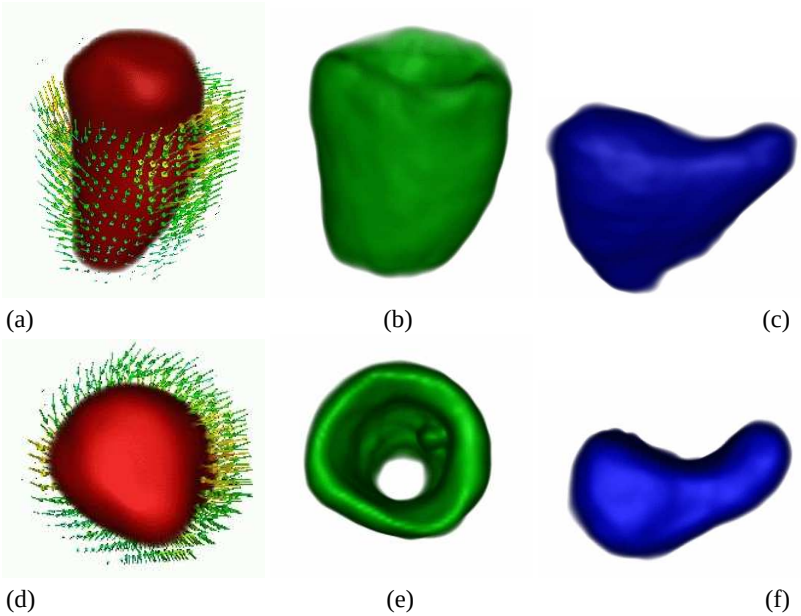


Figure 2. Examples of the volume renderings: (a), (c) the left ventricle with the deformation fields viewed from two different positions (b), (e) the myocardium viewed from the same two different positions and (c), (f) the right ventricle viewed from the same two different positions. Animations of the atlases are found at `http://www.doc.ic.ac.uk/~dp1/Conferences/MIUA04/ATLAS/.`

temporal alignment techniques, for example 4D deformable registration techniques, during the atlas construction. The use of a 4D deformable alignment technique will not only correct scaling, rotation and translation differences but also a certain amount in shape differences of the subjects allowing the atlas to provide better definition for the boundaries of each structure. In this case, the information regarding the shape variability will not be encoded in the tissue probability maps but in the deformation fields generated by the registration. Finally, techniques for atlas-based registration will be developed in order to allow the atlases to be used in the medical environment. The comparison of particular subjects with the atlases can assist the clinician to identify possible abnormalities in the subjects' cardiac anatomy and function.

References

1. American Heart Association. "Heart and stroke statistical update." http://www.americanheart.org/, 2002.
2. J. Mazziotta, A. Toga, A. Evans et al. "A probabilistic atlas of the human brain:theory and rationale for its development. The international consortium for brain mapping." *Neuroimage* **2**, pp. 89–101, 1995.
3. J. M. et al. "A four dimensional probabilistic atlas of the human brain." *Journal of the American Medical Informatics Association* **8(5)**, 2001.
4. A. Frangi, D. Rueckert, J. A. Schnabel et al. "Automatic construction of muliple-object three-dimensional statistical shape models: Application to cardiac modeling." In *In Information Processing in Medical Imaging: Proc 17th International Conference (IPMI'01)*, pp. 78–91. Davis, CA, 2001.
5. A. Frangi, D. Rueckert, J. A. Schnabel et al. "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling." *IEEE Transaction on Medical Imaging* **21(9)**, pp. 1151–1165, 2002.
6. J. Lötjönen, J. Koikkalained, D. Smutek et al. "Four-chamber 3D statistical shape model from cardiac short-axis MR images." In *Sixth Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'03)*, Lecture Notes in Computer Science, pp. 491–498. Springer-Verlag, 2003.
7. M. Lorenzo-Valdés, G. Sanchez-Ortiz, R. Mohiaddin et al. "Atlas based segmentation and tracking of 3D cardiac MR images using non rigid registration." In *Fith Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'02)*, Lecture Notes in Computer Science, pp. 642–650. Springer-Verlag, 2002.
8. D. Perperidis, A. Rao, M. Lorenzo-Valdés et al. "Spatio-temporal alignment of 4D cardiac MR images." In *Lecture Notes in Computer Science: Functional Imaging and Modeling of the heart, FIMH'03*. Springer, Lyon, France, June 5-6, 2003.
9. C. Studholme, D. Hill & D. Hawkes. "An overlap invariant entropy measure of 3D medical image alignment." *Pattern Recognition* **32**, pp. 71–86, 1999.
10. R. Chandrashekara, R. H. Mohiaddin & D. Rueckert. "Analysis of myocardial motion in tagged MR images using nonrigid image registration." In M. Sonka & J. M. Fitzpatrick (editors), *Proceedings of the SPIE International Symposium on Medical Imaging*, pp. 1168–1179. San Diego, California USA, 24–28 February 2002.
11. A. Rao, G. Sanchez-Ortiz, R. Chandrashekara et al. "Comparison of cardiac motion across subjects using non-rigid registration." In *Proceedings: Medical Image Computing Assisted Intervention, MICCAI 2002*. Tokyo, Japan, 2002.

Decoupling of Respiratory Motion with Wavelet and Principal Component Analysis

Fani Deligianni, Adrian Chung and Guang-Zhong Yang

Royal Society/Wolfson Foundation Image Computing Laboratory, Imperial College London, SW7 2BZ

Abstract. The establishment of patient specific models for minimal access surgical simulation requires the acquisition and co-registration of 2D endoscope/laparoscope video with 3D tomographic data with matched physiological status. The advent of *in vivo* catheter tip tracking devices offers the potential for improving the robustness and accuracy of current registration techniques in the presence of tissue deformation. For bronchoscope simulation, reliable extraction of respiratory motion allows retrospective gating of the acquired tracking data so that video bronchoscope views can be grouped according to different respiratory phases. In practice, the motion data recorded is coupled with patient and respiratory motion and the decoupling of the two is not trivial. This paper presents a novel motion decoupling technique for simplifying 2D/3D registration under the influence of normal respiratory motion. Wavelet analysis has been used to identify and remove episodes due to coughing and extreme breathing patterns. The technique has been validated with data acquired from 8 subjects, demonstrating the practical value of the proposed method.

1 Introduction

Minimal invasive surgery is increasingly being used in routine clinical practice as it significantly reduces patient trauma and recovery time. The technique, however, requires a high degree of manual dexterity and hand-eye coordination. The use of simulation devices has been proven to be an economical and time saving tool for acquiring, as well as assessing basic surgical skills. With the current systems, however, the lack of visual realism and tactile feedback represents major challenges to minimal invasive surgery tasks. To this end, computer simulation is becoming an important tool for acquiring, as well as assessing, basic surgical skills. In our previous study, we have developed a reliable 2D/3D registration technique, which incorporates structural matching, and inter-frame coherence for providing photorealistic rendering for bronchoscope simulation [1]. The method has shown great promise in establishing the camera pose during bronchoscope examination such that matched surface details can be extracted to augment the virtual bronchoscope views. The technique, however, assumes that there is no large tissue deformation between 3D tomographic data and 2D bronchoscope video. To simplify the 2D/3D registration process and accommodate general tissue deformation during examination, the use of catheter tip EM tracker provides a practical way forward. In a catheter tip tracking enabled bronchoscopy examination, a 6 DoF tracker is required to monitor the global position and orientation of the patient. This allows the cancellation of global motion, thus facilitating the localisation of the bronchoscope in relation to the CT scan volume. One of the issues involved in tracking the pose of bronchoscope camera is its movement due to respiratory motion. The trajectory acquired needs to be separated into different phases of the respiratory cycle such that the corresponding bronchoscope views can be co-registered with 3D tomographic data. Since during bronchoscope examination, both patient and respiratory motion affect the reading of the catheter tip, they need to be decoupled before further processing steps can be applied. The purpose of this paper is to provide a novel respiratory motion decomposition scheme to accurately monitor the respiratory cycle in the presence of large patient motion and filter out episodes due to coughing and extreme breath patterns.

2 Methods

2.1 Experimental Setup

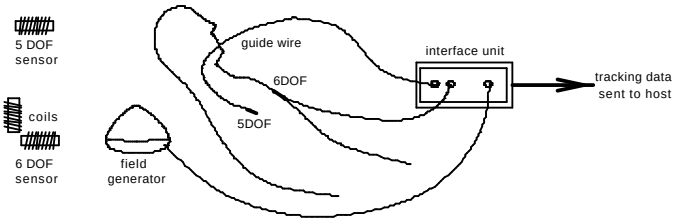


Fig 1: The basic configuration of using Aurora electromagnetic tracking systems in a clinical setting for removing global patient motion and recovering respiratory pattern.

The basic system layout is shown in Fig 1 where an Aurora (Northern Digital, Ontario, Canada) EM tracking system is used. The system provides both 5 DoF catheter tip and 6 DoF (MagTrax Reference) EM trackers. Additional positional trackers

{fd301,ajchung,gzy}@doc.ic.ac.uk

can be introduced to extract respiratory motion so that video bronchoscope views can be grouped according to different respiratory phases. The introduction of extra tracking devices, however, complicates the experimental procedure and may not be practical in clinical examinations. Due to the different frequency characteristics of the respiratory and global motion, it is possible to simultaneously acquire and decouple these motions by using a single sensor positioned on the chest wall. The decoupling of global and respiratory motion in this case, however, is not trivial as they are intertwined depending on the projection axes. Given the fact that the dominant cyclic variation of the signal is due to respiratory motion which has a constant principal direction of variation during a given time window, principal vector decomposition is used to extract the plane of respiratory motion over time, independent of global motion. To account for coughing and other extreme respiratory motions, wavelet analysis [2] was applied to detect and isolate these rapid motion episodes. In this paper, we present a new technique that uses principal component analysis based method to retrieve "hidden" respiratory patterns from position sensors in different axes and wavelet transform to filter out coughing and extreme breath patterns.

2.2 Decoupling of Respiratory Motion

Given that the dominant variation of the signal is due to respiratory motion and it has a constant direction during a given time window, then principal component analysis can be employed to decouple it. The principal vector is expected to lie on the same plane over time and, thus projecting the signal on this plane can isolate the signal variations due to respiration. We define as f the positional vector over time of the MagTrax reference sensor attached to the chest. Typically the respiratory pattern has a higher frequency than positional drift of the patient, a kernel box $d(t)$ is convolved to each axis of the motion sensor such that

$$\mathrm{f}^*(t) = f(t) - f(t) * d(t) \tag{1}$$

In discrete form, the samples over time can be represented as:

$$\mathrm{f}^*(t - w\cdot\Delta t),\quad \mathrm{f}^*(t - w\cdot\Delta t + \Delta t),\quad \ldots.\quad \mathrm{f}^*(t + w\cdot\Delta t) \tag{2}$$

A matrix h is defined for each time sample as:

$$h(t) = \left[\mathrm{f}^*(t - w\cdot\Delta t),\quad \mathrm{f}^*(t - w\cdot\Delta t + \Delta t),\quad \ldots.\quad \mathrm{f}^*(t + w\cdot\Delta t)\right]^T \tag{3}$$

h has dimension of $(2w+1)\cdot 3$, where 2w has a typical length of 30 sec. Therefore, the covariance matrix is produced as:

$$M(t) = \left(h(t) - \bar{h}(t)\right)\cdot\left(h(t) - \bar{h}(t)\right)^T \tag{4}$$

The principal vector $p(t)$ is estimated from the eigen decomposition of the covariance matrix. Finally, r is a scalar vector, which represents the decoupled respiratory motion. This is estimated as the projection of $\boldsymbol{f}^*$ on the principal vector.

$$r(t) = \mathrm{f}^*(t)\cdot p(t) \tag{5}$$

2.3 Wavelet Analysis

During bronchoscope examination, coughing involves major distortion to the airway. In order to identify these episodes wavelet analysis was used [2,3,4]. Wavelet analysis provides localised frequency analysis and has the potential to analyse signals that contain multiple non-stationary or transitory signal characteristics. Their main advantages are both its ability to perform local analysis and to preserve time information.

Let $y(t)$ be a function in the Hilbert space $L^2(R)$ of measurable, square-integrable one-dimensional functions with an average of zero, and denote $y_{2^j}(t) = 2^{-j} y\left(t/2^j\right)$. The wavelet transform of a function $f(t)$ at the scale 2^j and position t is given by the convolution product:

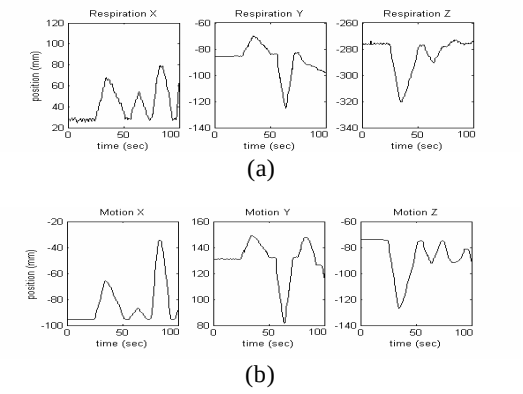
$$W_{2^j} f(t) = f * y_{2^j}(t) \tag{6}$$

The *dyadic wavelet transform* is the sequence of functions $\left(W_{2^j} f(t)\right)_{j\in Z}$, where Z represents the set of integers. In multiscale edge analysis, $y(t)$ is usually chosen to be the derivative of some smoothing function and, thus, the local maxima of $\left|W_{2^j} f(t)\right|$ indicate the positions where sharp signal variations occur. In order to identify these variations we define an energy function e with threshold x.

$$e(t) = \sum_{j<10} W_{2^j} f(t) > x \tag{7}$$

3 Validation

Eight subjects were recruited for this study with their respiration monitored by an Aurora (Northern Digital, Ontario, Canada) 6 DoF EM tracker positioned on the sternum. Their respiration was monitored for about two minutes. Global motion was introduced to the couch on which subjects were seated to simulate the sensor reading coupled with two types of motion. During the experiment the relative position of the subject and the couch was fixed and the subjects were allowed to have extreme respiratory movements including coughing. An additional EM tracker was placed on the base of the couch so that its relative movement to that of the sensor located on the subject's chest wall provided reference readings purely due to respiration. This information was then used to validate the accuracy of the motion decoupling results based on the single sensor method. The results of this study also demonstrate that episodes related to coughing can be reliably identified and filtered out. The subject was asked to cough at a particular time, while slight global motion was introduced. An example is demonstrated in Fig 3 and has been used to validate the wavelet technique.



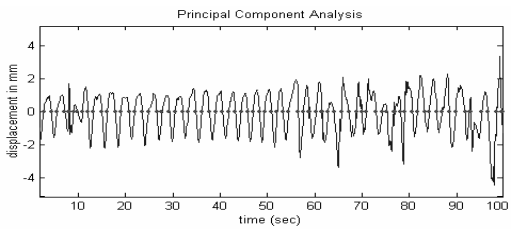


Fig 2: Example position traces sampled by the EM tracker and the extracted respiratory motion pattern. (a) *x, y, z* positional component of the data received from the EM tracker attached on the chest of one of the subjects, (b) *x, y, z* positional component of the data received from the EM tracker attached on the couch, (c) The recovered respiratory component.

4 Results

Fig 2 demonstrates typical results derived from the proposed respiratory motion decoupling technique. It illustrates the *x, y, z* positional component of the data received from the EM tracker attached on the chest of one of the subjects (a), as well as the *x, y, z* positional component of the data received from the EM tracker attached to the base of the couch, (b). Fig 2(c) represents the recovered respiratory component after applying the proposed technique described in section 2.1 to the data received from the EM tracker on the subject's chest. In order to validate our technique the correlation between the recovered waveform (section 1.1) and the waveform estimated by applying principal component analysis to the vector between the two EM tracker tool tips has been calculated for eight subjects. Table 1 outlines the correlation accuracy of the results derived from the 8 subjects studied, which achieved an overall accuracy of 82.7%. This indicates the ability of the technique in extracting the respiratory signal in spite of the fact that it is not apparent in the initial EM tracking data due to global motion. Fig 3 demonstrates the wavelet analysis of the X, Y and Z positional data that have been acquired from an EM tracker attached on the skin of a healthy subject. The wavelet coefficients are plotted for each signal component. Fig 3(d) represents the energy function of the X-wavelet coefficient graph that has been defined within equation (7). The time interval where the patient coughs shows up clearly as a peak that corresponds to 88 sec. By using the wavelet decomposition scheme, extreme breathing patterns can be reliably identified.

Subjects	1	2	3	4	5	6	7	8
Correlation	0.918	0.962	0.805	0.794	0.879	0.791	0.786	0.687

Table 1: Correlation of the waveform estimated by the suggested respiratory decoupling motion technique with the waveform estimated by relative position of two sensors.

5 Discussion and Conclusions

In conclusion, we have presented a novel method of decoupling respiratory motion from the signal received from an EM tracker attached to the chest. We used wavelet analysis to filter out episodes due to coughing and other extreme breathing patterns. Principal component analysis has been employed to decompose the respiratory cycle from global motion. It should

be noted that the basic assumption used for respiratory motion decoupling is that the global positional drift has a relatively low frequency in comparison to respiratory cycles. When there is sudden motion involving large acceleration, it is likely that rapid changes in sensor readings will be introduced, in this case, the algorithm based on principal component analysis will lead to detection errors. Since we have used wavelet analysis to sense rapid sensor movement, these events will also be isolated along with coughing and extreme respiratory motion. The results in this study indicate that the respiratory motion component can be reliably extracted. Even though the respiratory signal has non-zero components in all *x*-*y*-*z* axes and it is blended with global motion, they can be reliably decoupled with the proposed technique. Furthermore, episodes related to coughing can be detected easily and filtered out using wavelet analysis, which is appropriate for detecting localised signal variations. The technique developed can facilitate a 2D/3D registration regime under deformation and increase the accuracy of the registration process.

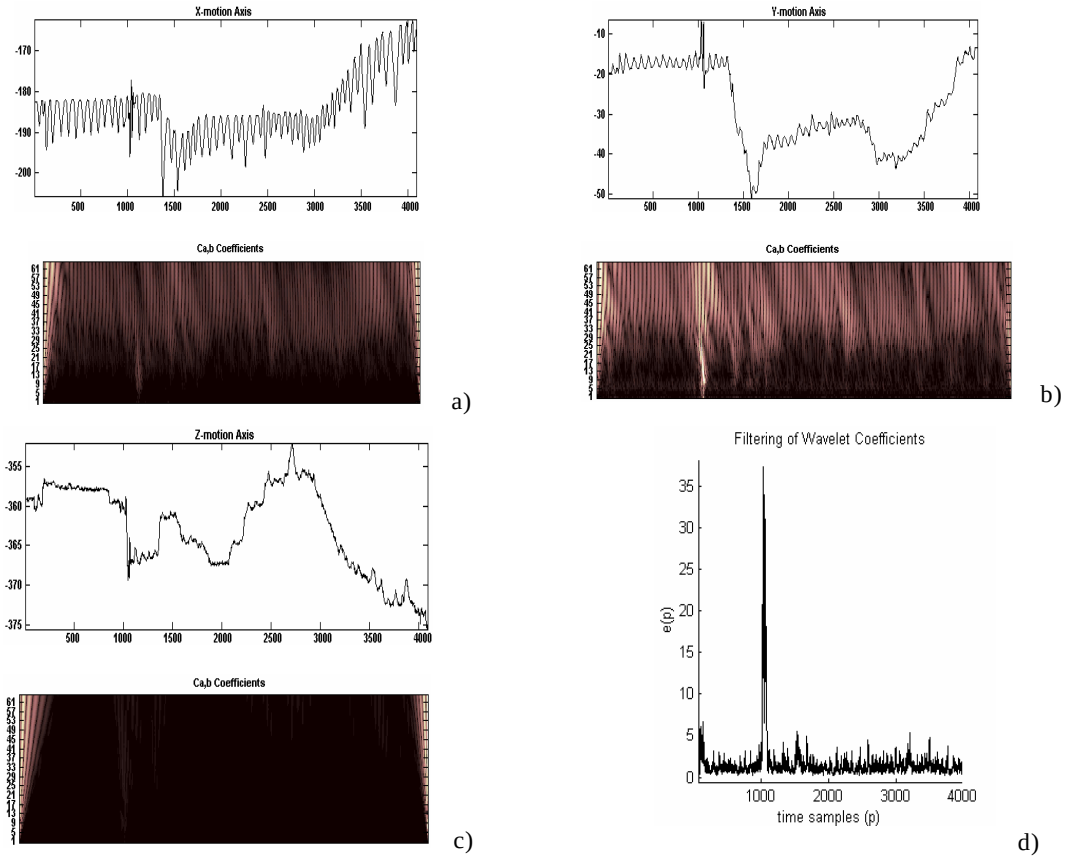


Fig 3: Wavelet analysis: X, Y and Z positional data has been acquired from an EM tracker attached on the skin of a healthy subject in order to assess the wavelet technique of filtering sudden movements similar to those that are introduced in a bronchoscopy session when the patient coughs. We used Daubechies (db10) mother wavelets to analyze the signals. a-b-c demonstrate the three positional (x,y,z) signals and the correspondent wavelet coefficients. e) Demonstrates the energy function after analyzing the X-motion signal. Similar results have been acquired from analyzing either the Y or the Z component.

References

1. Deligianni F., Chung A., Yang G. Z.: pq-Space 2D/3D Registration for Endoscope Tracking. Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI03), 1, (2003), 311-318
2. Nikolopoulos S., Alexandridi A., Nikolakeas S., Manis G.: Experimental analysis of heart rate variability of long-recording electrocardiograms in normal subjects and patients with coronary artery disease and normal left ventricular function. Journal of Biomedical Informatics, 36(3), October (2003), 202-217
3. Figliola A., Rosso O.A., Serrano E.: Detection of delay time between the alterations of cardiac rhythm and periodic breathing. Physica A: Statistical Mechanics and its Applications, 327(1), July (2003), 174-179
4. Senhadji L., Carrault G., Bellanger J.J., Passariello G.: Comparing wavelet transforms for recognizing cardiac patterns. IEEE Engineering in Medicine and Biology Magazine, 14(2), March-April (1995), 167-173

Image-based reduction of artefacts in radial imaging

K J Lee[a]*, M N Paley[a], P D Griffiths[a] and J M Wild[a]

[a]Academic Radiology, University of Sheffield, Sheffield, UK

Abstract. Previously, we have described the method of generalised projections (MGP) to correct for phase discontinuities in single- and multi-shot echo-planar imaging (EPI). We investigate through computer simulations and phantom experiments how well the method works for phase inconsistencies in radial trajectories. Two models of phase error were investigated: random isotropic shifts such as might be found in multi-echo radial EPI, and anisotropic shifts such as might be caused by gradient delays in any radial sequence. With isotropic shifts, we found that MGP can reduce severe artefacts, but it is limited by the fact that it is not guaranteed to converge, and the method is not recommended for low artefact levels. The method was found to fail for anisotropic shifts.

1 Introduction

There has been renewed interest in acquiring k-space data using radial trajectories (projection reconstruction) as an alternative to Cartesian trajectories, due to their relative insensitivity to patient motion [1] and their suitability for fast imaging [2]. However, miscentering of radial k-space lines can cause streak artefacts that are significant because the phase inconsistencies affect the centre of k-space. Miscentering caused by gradient timing delays may be corrected by adjustment of gradients after suitable measurement of the delays [3, 4]. In projection reconstruction with multi-gradient echoes (radial EPI), T_2^* weighting and an echo number dependent constant phase term generate streaking artefacts [5], as do differences in timing between positive and negative readouts which lead to misalignment between echoes [6]. These effects can be corrected by measurement of anti-parallel projections, but this requires double the number of projections for a particular resolution [5].

In this paper we examine an image-based, post-processing method of reducing artefacts caused by phase and echo shifts terms. The aim is to correct for miscentering without the additional measurements required in other methods. This is equivalent to finding an image-based method using data collected over 180° only, as data acquired over the second half of k-space constitutes anti-parallel projections which may be used for correction. The algorithm is based on the method of generalised projections (MGP) which is the extension of the well-known projections onto convex sets (POCS) algorithm to non-convex sets.

2 Description of algorithm

The proposed algorithm is an extension to radial trajectories of an algorithm that we have used for ghost reduction in Cartesian EPI [7–9]. It is also similar to MGP algorithms previously used for object motion correction [10, 11]. Our algorithm differs from that for object motion correction by phase correcting in sinogram space, instead of k-space.

The algorithm starts by reconstructing the artefacted image from the original echoes $S(k_r, k_\theta)$ by regridding on to a Cartesian grid followed by a Fourier transform [12]. Next, we impose an image constraint by identifying a region of support (ROS) around the parent image outside which any signal is known to be artefactual. The region outside is set to zero, and the masked image is inverse Fourier transformed, and sampled along the designed trajectories to give a set of "reference echoes", $S_{ref}(k_r, k_\theta)$. We assume the artefacts in the image are due to constant phase terms and echo shifts which are angle dependent. By the Fourier shift theorem, this equates to zero and first-order phase shifts in the sinogram. Constraints in sinogram space are now imposed as follows: the phase of the sinogram from $S_{ref}(k_r, k_\theta)$, up to first-order, is assumed to be more correct than that of $S(k_r, k_\theta)$, and replaces it. The zero and first-order phase corrections are obtained by linear fitting of the phase difference between reference and original sinograms. This completes one iteration of the algorithm, which is then iterated until an endpoint. Setting the phase as above is a nonconvex constraint [8, 10], and therefore convergence is not guaranteed [13], and iterations may need to be terminated manually.

*k.j.lee@sheffield.ac.uk

3 Method

3.1 Computer simulations

To investigate the algorithm simulations were performed on a test phantom image in MATLAB. The image size was 256×256 pixels, consisting of two asymmetrically placed, overlapping discs, one real and one complex (Fig. 1). The image was inverse Fourier transformed, giving a complex k-space with an asymmetrical peak. From this, corrupted k-spaces are obtained by sampling over 180° at 128 angles, along *incorrect* trajectories according to two models, simulating the errors arising in multi-gradient echo acquisition and gradient delays respectively.

Let the errors in each view be described by a constant phase term ϕ and a total translation error in k-space trajectory $\triangle\mathbf{k}$. Our first model describes errors arising from different eddy current behaviour between positive and negative readouts in a multi-gradient echo acquisition. Here we assume that this leads to random constant phase terms ϕ and radial echo shifts $\triangle k_\parallel$ (the portion of $\triangle\mathbf{k}$ parallel to the readout direction) [5, 6]. Thus we ignore $\triangle k_\perp$, the component of $\triangle\mathbf{k}$ perpendicular to the readout direction, and the shifts are therefore *isotropic*. The phase and echo shifts applied were varied from a maximum of ± 0.125 radians,± 0.5 time points to test the performance of the algorithm at different artefact levels. MGP was then applied to estimate and correct for $\phi(\theta)$ and $\triangle k_\parallel(\theta)$ directly. The algorithm was stopped when the change in $\triangle k_\parallel$, $\delta(\triangle k_\parallel) < 0.1$ time points.

The second model is known to describe miscentered trajectories due to gradient delays [3, 4]. In this gradient delay model, $\triangle\mathbf{k}$ is given by a vector sum of the separate translation errors in each physical gradient: $\triangle\mathbf{k} = \triangle k_x p_x \mathbf{i} + \triangle k_y p_y \mathbf{j}$, where $p_x, p_y \in [0...1]$ are the relative gradient strengths. Similarly, the constant phase error is given by the sum of separate constant errors in each gradient: $\phi = \phi_x p_x + \phi_y p_y$. The shifts can also be described by [3]:

$$\triangle k_\parallel(\theta) \propto \triangle k_x \cos^2\theta + \triangle k_y \sin^2\theta \tag{1}$$

$$\triangle k_\perp(\theta) \propto \sin\theta\cos\theta\,(-\triangle k_x + \triangle k_y) \tag{2}$$

Note that in this case of non-zero $\triangle k_\perp$, the shifts are *anisotropic*. Because our proposed algorithm only estimates shifts in the *radial* direction, it can not directly correct for $\triangle k_\perp$. However, we use the algorithm to estimate values of $\triangle k_\parallel(\theta)$, and a least-squares fit with Eq. (1) to estimate the constants $\triangle k_x$ and $\triangle k_y$. These can then be used to calculate $\triangle k_\perp(\theta)$ using Eq. (2). Delays were simulated of $\triangle k_x, k_y = [-0.5 \ldots 0.5]$ time points. Phase differences were simulated up to a maximum of $\phi_x, \phi_y = +0.1$. Iterations were stopped when the change in $\triangle k_x, k_y < 0.01$ time points or after 20 iterations.

Artefact levels were measured by M, the mean absolute difference between the pixels in artefacted and original images over the ROS.

3.2 Phantom experiment

Phantom data was acquired on a 1.5T MR scanner (Eclipse, Philips Medical Systems), using an in-house developed multi-echo radial sequence [14]. A sequence plot is shown in Fig. 2. Note that this gives two echoes at every view. 128 views were collected over 180° with $TE_1/TE_2 = 3.5/5.5$ ms. By using the first echoes for odd views and second echoes for even views (taking care of time reversal), we simulate an interleaved dual-echo radial EPI acquisition. By using only the first echoes for all views, a reference image can be created for comparison that should be free from misalignment artefacts caused by gradient reversal. Iterations were stopped when the change in $\triangle k_\parallel < 0.1$ time points.

4 Results

4.1 Computer simulations

With isotropic shifts, the algorithm always converged after ≤ 2 iterations. Fig. 3 shows the mean and standard deviation values of M for 10 sets of randomly corrupted k-spaces before and after correction at different starting artefact levels. We see that the algorithm can reduce the artefact level (Fig. 4) but there is a cross-over point beyond which the algorithm degrades the image. The explanation for this lies in Fig. 5, which shows the final estimated values of $\triangle k_\parallel(\theta)$ for a starting $M = 0\%$. Although the algorithm has estimated the values to within 0.05 time points of the correct values (zero), this is sufficient error to degrade the image slightly, giving a final

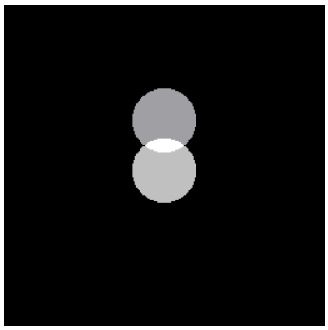


Figure 1. Magnitude image of test object. The lower disc is real and the upper disc complex

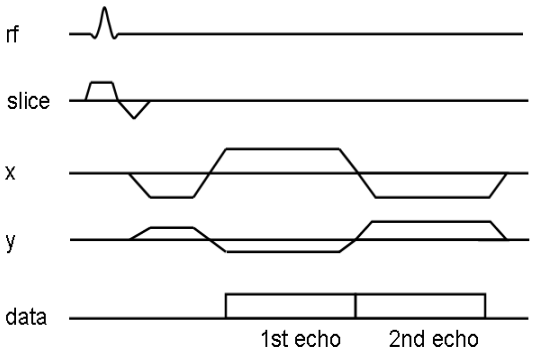


Figure 2. Multi-echo radial pulse sequence

$M = 2.4\%$. Unfortunately, changing the end-point criterion in an attempt to decrease the error, for example by choosing $\delta(\triangle k_{\|}) < 0.01$ instead of 0.1, makes the algorithm unstable and divergent (not shown). The reason for this is because the algorithm does not project onto a convex set and therefore convergence is not guaranteed.

With anisotropic shifts, for data acquired over 180°, the algorithm never converged and had to be stopped after 20 iterations, the error in $\triangle k_{\|}(\theta)$ being always too large to estimate $\triangle k_x$ and $\triangle k_y$ with sufficient accuracy for convergence. Fig. 6 shows the values of M for anisotropically corrupted k-space before and after correction at different starting artefact levels.

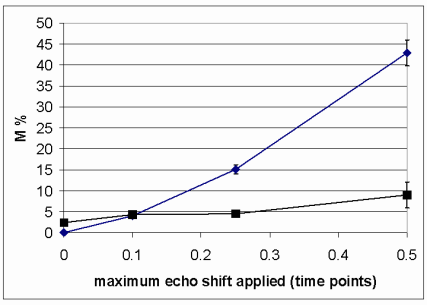
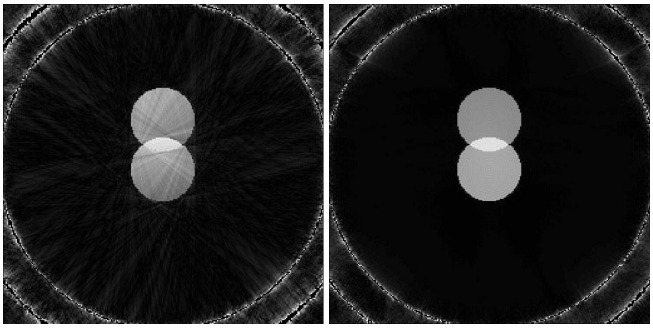


Figure 3. M vs applied artefact level for isotropic shifts (♦ before correction ■ after correction)



(a) $M = 40.0\%$ (b) $M = 9.5\%$

Figure 4. Isotropically corrupted images (a) before (b) after MGP correction

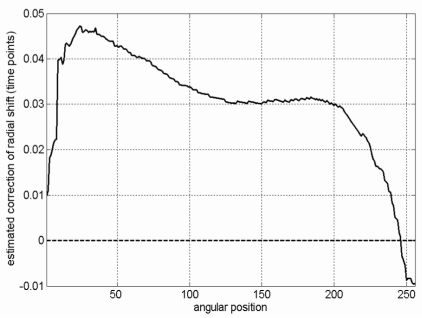


Figure 5. Estimate of $\triangle k_{\|}$ for image with starting $M = 0\%$.

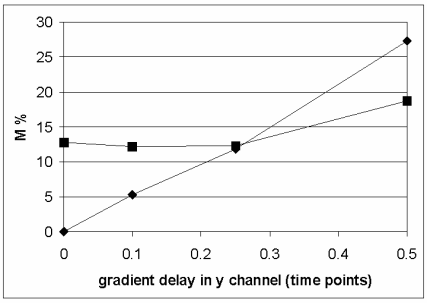
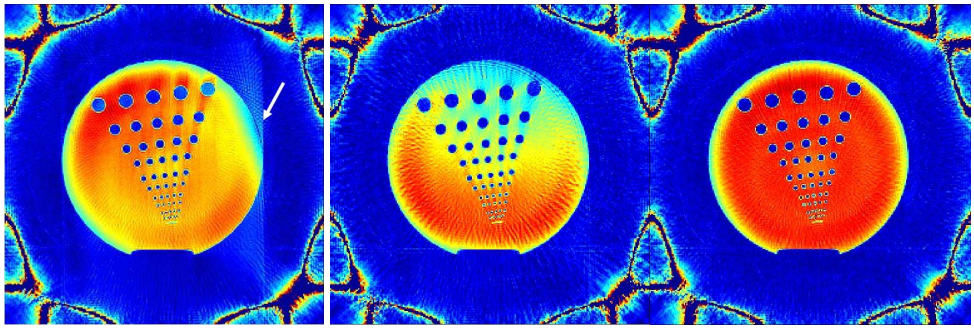


Figure 6. M vs applied artefact level for isotropic shifts (♦ before correction ■ after correction)

4.2 Phantom experiment

The algorithm converged after 2 iterations. Figure 7 shows the reference image reconstructed from first echoes only, the simulated dual-echo radial EPI, and the MGP corrected image respectively. There are some artefacts in the reference image (arrow), presumably from hardware instability. The simulated radial EPI shows severe shading



(a) (b) (c)

Figure 7. Real phantom (a) original test image (b) shaded artefacts from echo misalignments (c) MGP corrected.

artefacts from echo misalignment. These are removed by the MGP correction, which also restores uniformity.

5 Discussion and Conclusions

In previous work, we found that MGP is effective only if we impose constraints through some *a priori* model of phase variation in order to reduce the degrees of freedom. We had hoped that in radial imaging, with its inherent oversampling, MGP might be more robust here than with Cartesian trajectories. Our simulations have shown that while good correction is possible where shifts are isotropic, by not having a guarantee of convergence, the algorithm is limited to correcting large artefacts only. An example of this was shown by our phantom experiment. With low artefact levels, MGP cannot reduce the error in its estimates by reducing step size, as this led to divergence. Therefore, where artefact levels are low, the correction is not recommended. MGP correction has also been shown to fail where the shifts are anisotropic.

References

1. G. Glover & J. Pauly. "Projection reconstruction techniques for reduction of motion effects in MRI." *Magn Reson.Med.* **28(2)**, pp. 275–289, 1992.
2. C. Bergin, J. Pauly & A. Macovski. "Lung parenchyma: projection reconstruction MR imaging." *Radiology* **179(3)**, pp. 777–781, 1991.
3. D. Peters, J. Derbyshire & E. McVeigh. "Centering the projection reconstruction trajectory: reducing gradient delay errors." *Magn Reson.Med.* **50(1)**, pp. 1–6, 2003.
4. O. Wieben, E. Brodsky, C. Mistretta et al. "Correction of trajectory errors in radial acquistion." In *Proceedings of 11th Annual Meeting of International Society for Magnetic Resonance in Medicine*, p. 298. 2003.
5. V. Rasche, D. Holz & R. Proksa. "MR fluoroscopy using projection reconstruction multi-gradient-echo (prMGE) MRI." *Magn Reson.Med.* **42(2)**, pp. 324–334, 1999.
6. T. Schaeffter, S. Weiss, H. Eggers et al. "Self-referencing multi-gradient-echo projection reconstruction." In *Proceedings of 11th Scientific Meeting of International Society for Magnetic Resonance in Medicine*, p. 207. Toronto, 2003.
7. K. Lee, D. Barber, M. Paley et al. "Image-based EPI ghost correction using an algorithm based on projection onto convex sets (POCS)." *Magn Reson Med.* **47(4)**, pp. 812–817, 2002.
8. K. J. Lee, N. G. Papadakis, D. C. Barber et al. "A method of generalized projections (MGP) ghost correction algorithm for interleaved EPI." *IEEE Transactions on Medical Imaging* **23(7)**, pp. 839–848, 2004.
9. K. J. Lee, M. Paley, D. C. Barber et al. "Image-based ghost reduction of amplitude discontinuities in k-space with projections onto convex sets (POCS)." In *Proceedings of 12th Scientific Meeting of International Society for Magnetic Resonance in Medicine*, p. 2198. 2004.
10. M. Hedley, Y. Hong & D. Rosenfeld. "Motion artifact correction in MRI using generalized projections." *IEEE Transactions on Medical Imaging* **10(1)**, pp. 40–46, 1991.
11. T. Trouard, Y. Sabharwal, M. Altbach et al. "Analysis and comparison of motion-correction techniques in diffusion-weighted imaging." *J Magn Reson Imaging* **6(6)**, pp. 925–935, 1996.
12. J. Jackson, C. Meyer, D. Nishimura et al. "Selection of a convolution function for Fourier inversion using gridding." *IEEE Transactions on Medical Imaging* **10(3)**, pp. 473–478, 1991.
13. A. Levi & H. Stark. "Image-restoration by the method of generalized projections with application to restoration from magnitude." *Journal of the Optical Society of America A-Optics Image Science and Vision* **1(9)**, pp. 932–943, 1984.
14. J. Wild, M. Paley, L. Kasuboski et al. "Dynamic radial projection MRI of inhaled hyperpolarized 3He gas." *Magn Reson.Med.* **49(6)**, pp. 991–997, 2003.

Noise Filtering and Testing for MR Using a Multi-Dimensional Partial Volume Model

M. Pokrić*, N.A. Thacker, D.C. Williamson.

Imaging Science and Biomedical Engineering
School of Medicine
University of Manchester
Manchester M13 9PT, UK

Abstract.
One of the most common problems in image analysis is the estimation and removal of noise or other artefacts using spatial filters. Common techniques include Gaussian Filtering, Median Filtering and Anisotropic Filtering. Though these techniques are quite common they must be used with great care on medical data, as it is very easy to introduce artifact into images due to spatial smoothing. The use of such techniques is further restricted by the absence of a 'gold standard' data against which to test the behaviour of the filters. Following a general discussion of the equivalence of filtering techniques to likelihood based estimation using an assumed model, this paper describes an approach to noise filtering in multi-dimensional data using a partial volume data density model. The resulting data sets can then be taken as a gold standard data for spatial filtering techniques which use the information from single images. We demonstrate equivalence between the results from this analysis and techniques for performance characterisation which do not require a 'gold standard'.

1 Introduction

Noise filtering on any data involves the assumption of a specific image generation mechanism or image model. The process of Gaussian smoothing can be interpreted as consistent with a likelihood estimation of the central value within a region. This is done on the assumption that the data within this region can be described by some functional model with the expected grey level residuals being drawn from a Gaussian noise distribution with variance inversely proportional to the spatial Gaussian weighting term. The specific form of the assumed model is best understood by considering a more simplistic problem first. Gaussian filtering removes high spatial frequency content from the structure of images. A less destructive approach to noise filtering is based on the concept of anisotropic filtering, where the data is preferentially smoothed along a direction selected in order to minimise the loss of spatial structure in the image. One particular variant of this we call *Tangential Filtering*. Here, the tangential direction to the local image slope is computed and the data is smoothed by taking the weighted average of points along this line situated on either side of the central value. It is relatively straight forward to see that averaging of multiple values in this way assumes that the data can locally be fitted to a 1D line and selection of the tangential direction results in the least destructive impact on edge structure. In Gaussian filtering, we can interpret this process as an averaging of multiple estimations of the central value for any pair of pixels with equal weight in the Gaussian kernel. The class of functions for which this would be an appropriate model would include Cartesian polynomials (expanded as a function of shifts (x, y)) from the central location (x_0, y_0), either with no even terms, or at least exact cancellation of the magnitudes of even power coefficients. For example:

$$I(x - x_0, y - y_0) \;=\; a \;+\; bx \;+\; cy \;+\; dxy \;+\; ex^2 \;-\; ey^2 \;+\; ...$$

Though the only global function for which the model would work correctly at every image location would be an inclined plane. When described in this way it is easy to see the over-simplicity of this model in comparison to the structures found in real medical images. This produces the characteristic problems with Gaussian filtering manifested by image smoothing and the loss of sharp edge structures.

As a direct contrast to spatial image filtering, which assumes specific forms of spatial correlation between grey level values, multi-dimensional tissue segmentation algorithms rely instead upon corellations between multiple measurements of the same physical location (voxel) using different imaging modalities. The typical approach involves building a model not of spatial structure but of grey level density distribution. A common form of noise removal which makes use of grey-level density distribution is the median filter, which can be considered as a bootstrapped likelihood estimator.

*Email: maja.pokric@man.ac.uk

2 Method Overview

We will assess the stability of the selected filtering schemes first by using a Monte-Carlo approach. Here a small quantity of noise is added to the input image and the relative change in output grey-level values is measured. However, this technique is not sufficient as an evaluation, as it will not measure failure of the implicit filtering model. By interpreting each noise filtering technique as an estimation process we suggest that it is possible to assess the validity of the filtering method by observing the number of grey-level values which have changed by more than 3 standard deviations of the estimated image noise, i.e. a Residual Outlier Measurement (ROM). To estimate the image noise to set the scale of this measurement, we use a technique we call Local Noise Estimation (LNE). This is based upon the observation that high order derivatives are heavily corrupted by image noise. On the assumption of uniform independent noise we can use error propagation to predict the expected level of noise on any order of spatial derivative. By measuring the width of the distribution of derivatives around the peak at zero we can get an estimate of the original image noise by scaling with the appropriate error propagation factor. We use second derivatives which means that the technique is also consistent with measuring the deviation from a purely linear spatial grey-level model. Although this method will estimate the image noise well (within a few percent) for data with spatially uncorrelated noise, we cannot use this technique to estimate noise following spatial filtering. The Monte-Carlo assessment and the ROM measure are complementary aspects of performance, the first giving an indication of the best case noise filter on the assumption that the filter model is adequate and the second giving an indication of how often this model is inappropriate. Importantly, neither technique requires a "gold standard" to test against. In order to illustrate that these techniques give a complete characterisation of the algorithm performance we will need a surrogate "gold standard". We choose to use for this the Multi-Spectral Filtering method.

The Multi-Spectral Filtering method is based on multi-spectral image analysis which takes into account the effect of partial voluming [1, 2]. Parameters of the model are iteratively estimated by maximising the likelihood of the data distribution using the Expectation Maximisation (EM) approach [1, 3, 4]. This involves calculating the most likely volumetric contribution of a tissue to the voxel, using the conditional probabilities of all possible tissue volume generation mechanisms n (either a pure or mixture of tissues) $P(n|\mathbf{g})$, rather than simply calculating the conditional probability of the tissue classes. Noise free estimates of the individual image grey-levels g'_i can be calculated using

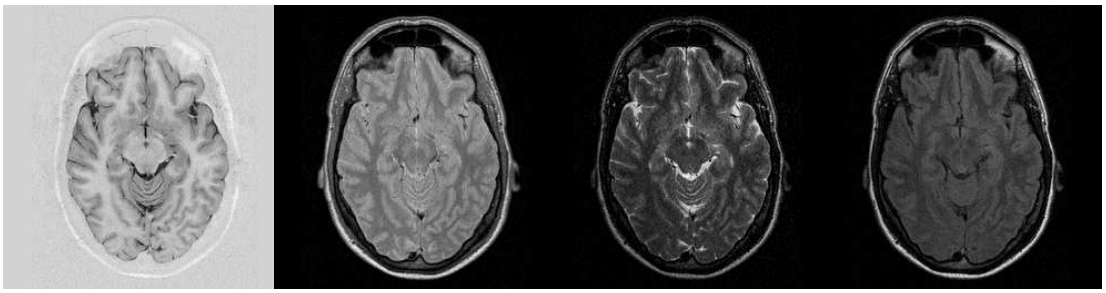
$$g'_i \;=\; g_i \; P(O|\mathbf{g}) \;+\; \sum_t g_{it} \; P(t|\mathbf{g}) + \sum_t \sum_s g_{it} \; P(ts|\mathbf{g})$$

Where g_{it} is the expected pure tissue grey level for image i, O is outlier class, t is pure tissue class, and ts is a class for mixture of tissues t and s. This formulation implicitly reverts to the original image grey-level for outlier data (large $P(O|\mathbf{g})$). In addition, failure to model individual voxels can be identified by checking for consistency between the filtered image and original. In particular, any reconstructed grey-level value which differs from the original by more than 3 standard deviation of the image noise can be said to be inconsistent with the model. If necessary this value can then be replaced with the original value in order to preserve all significant information present in the original image.

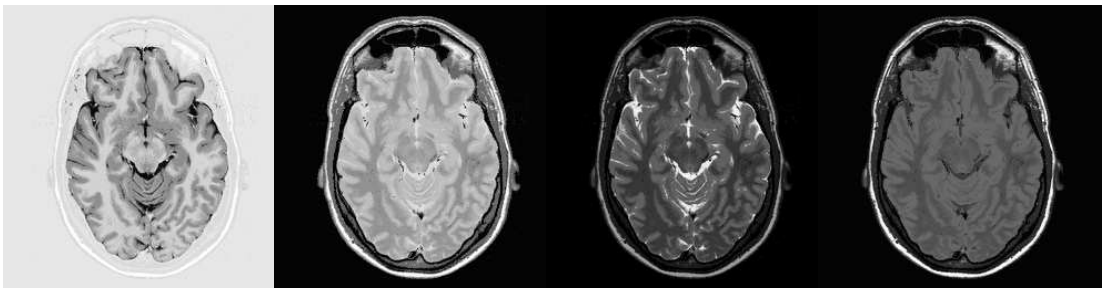
In order to use the results from the multi-spectral technique as gold standard there are a few issues which must be addressed. The first is that multi-spectral filtering is expected to eliminate spatially distributed errors, such as field inhomogeneity. Although these are expected to be small in our data (in comparison to intrinsic image noise) they also vary between acquisitions and will bias comparisons of residual distributions. In order to eliminate the majority of these effects we construct our pseudo-gold standard by removing a smooth estimate of local difference between the multi-spectral reconstruction and the original image, (constructed using a 5 pixel standard deviation (S.D.) Gaussian kernel). Secondly, the multi-spectral noise filtering process removes noise in a tissue dependant manner (also removing genuine tissue variability which cannot be interpreted as a partial volume process), so that residual difference distributions are not simple Gaussians. Therefore we fit all residual distributions within 3 S.D. of the estimated image noise with the sum of two Gaussians with the same mean but individual widths and normalisations (A and B). It is the results of this fitting process which we wish to reconcile with the results from Monte-Carlo and ROM quantification.

3 Results

Original and reconstructed images following co-registration and partial volume analysis are shown in Figure 1.
The noise level ($\sigma_{original}$) in each image was estimated using the LNE technique. Tangential filtering was applied by averaging over three pixels (one central and two either side). Gaussian filtering was for spatial filter with S.D.



(a) Inversion Recovery (TSE) (b) Variable Echo (PD) (c) Variable Echo (T2) (d) FLAIR



(e) Filtered IRTSE (f) Filtered VE (PD) (g) Filtered VE (T2) (h) Filtered FLAIR

Figure 1. Image sequences and partial volume filtered counterparts

of 1 pixel. Median filtering was over the local neighbourhood of 9 pixels. The Monte-Carlo stability analysis estimates of the fraction of noise remaining after filtering and the number of voxels lying beyond 3 S.D. of the model (ROM) are given in Table 1.

	LNE	Median Filtering	Gaussian Smoothing	Tangential Smoothing	Multi-Spectral
IRTSE	58.76	0.64 (1559)	0.27 (2405)	0.66 (698)	0.22 (1689)
VE(PD)	64.06	0.66 (1466)	0.26 (3127)	0.68 (530)	0.20 (1804)
VE(T2)	58.2	0.63 (1287)	0.26 (1909)	0.69 (426)	0.17 (938)
FLAIR	52.4	0.63 (1934)	0.27 (3827)	0.69 (966)	0.13 (4971)

Table 1. Monte-Carlo and ROM estimates

Following partial volume filtering the reconstructed images were corrected for low frequency spatial noise as described in the methods section. The resulting images were then used as the basis for evaluation of spatial filtering techniques applied to the original images (Table 2).

	Original					Median Filter				
	Mean	A	σ_1	B	σ_2	Mean	A	σ_1	B	σ_2
IRTSE	-4.6	0.47	44.5	0.53	75.3	0.39	0.39	21.9	0.61	63.7
VE(PD)	-4.9	0.54	37.5	0.46	90.6	-7.7	0.60	25.7	0.40	81.5
VE(T2)	-7.0	0.34	23.9	0.66	62.9	-7.1	0.58	21.7	0.44	61.5
FLAIR	0.27	0.50	37.1	0.50	93.9	0.15	0.49	20.7	0.51	74.0
	Gaussian Smoothing					Tangential Smoothing				
	Mean	A	σ_1	B	σ_2	Mean	A	σ_1	B	σ_2
IRTSE	1.1	0.35	14.5	0.65	65.3	-1.1	0.46	29.3	0.54	65.4
VE(PD)	0.77	0.36	10.6	0.64	52.2	4.3	0.51	24.2	0.49	75.1
VE(T2)	1.12	0.35	8.6	0.65	43.7	4.4	0.41	17.2	0.59	49.7
FLAIR	0.79	0.38	10.6	0.62	56.2	0.2	0.43	19.7	0.57	68.1

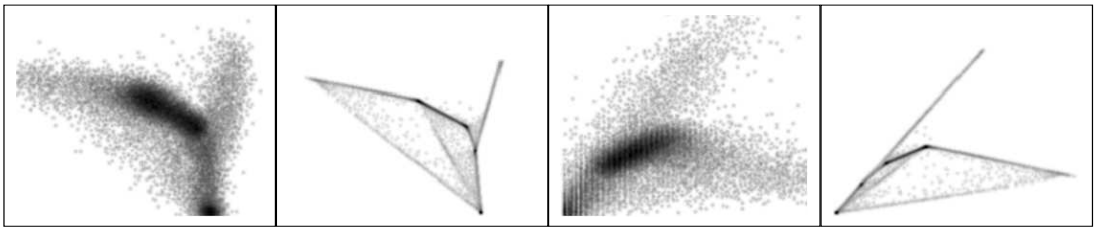
Table 2. Quantitative Performance of Spatial Filtering

Though the fits to the double Gaussian are less stable than those from the Monte-Carlo the results confirm that for the central narrow component the widths of the residual distribution vary by amounts consistent with the prediction. In addition the proportion of the secondary Gaussian components and secondary peak widths give an estimate of the number of failures. While the Gaussian filter reduces the width of both the primary and the secondary Gaussian distribution, the proportion of data in the broader part of the distribution has also increased. The proportion of data

in the secondary Gaussian remains the same for tangential smoothing and median filtering as for the original image. However, there is additional broadening of the secondary distribution for the median filter, as predicted by the ROM data. The data are in general reconcilable with ROM estimates, with the one exception being the lack of evidence for significant outlier proportion in multi-spectral filtering of FLAIR images. We believe that this is most likely due to flow artefact in the FLAIR images which has then been removed from the reference image by low frequency residual subtraction.

4 Discussion and Conclusions

Multi-spectral filtering can be considered as a regression onto the lines joining pairs of pure tissue locations in the multi-dimensional grey level space, followed by a weighting with pure tissue values according to the Bayesian priors. Any voxels composed of pure tissues of appropriate mean values will therefore have the noise on each grey level removed in such a way as to make the grey level value more consistent with the estimated position along this partial volume line (Figure 2). The results of such an analysis can be also interpreted as a method of data fusion, where data from alternative modalities are combined in order to improve the data from each. The results demonstrate that such an approach does not produce the loss of high spatial frequency structure inherent in even the most careful spatial filtering schemes.



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

Figure 2. Grey Level Distributions before (a) and after (b) Partial Volume Noise Filtering for IRTSE and VE(PD) images, and before (c) and after (d) Partial Volume Noise Filtering for VE(T2) and FLAIR images.

There is a subtle but important difference between using models based upon spatial distribution and those based upon partial volume behaviour for MR analysis. The former can only be determined from example data and there can never be a spatial model which will be appropriate for the contents of all biological images. Grey level density models however, have statistical characteristics which are purely determined by the acquisition process (i.e. the underlying physics of the measurement process). We might therefore expect that if we knew enough about the image formation process for a particular imaging protocol we may be able to construct a model which is true for all images from a particular acquisition containing equivalent tissue types. This approach to filtering may therefore be regarded as a gold standard for testing of spatial filtering techniques. In general multi-spectral filtering is no more destructive to image contents than tangential smoothing or median filtering but removes more image noise than Gaussian smoothing. We therefore suggest that such image reconstruction techniques may be a useful way of getting the most from multiple MR acquisitions. This paper has shown how techniques generally used for tissue segmentation can be used to provide noise filtering of multi-spectral images based upon an analysis of partial volume structure. We also conclude that the combination of an ROM and a Monte-Carlo stability analysis are sufficient to predict the important characteristics of a noise filtering scheme. Software and test data are available on the web (www.tina-vision.net).

Acknowledgements

This work was part of EC funded project "Performance Characterisation in Computer Vision" IST-1999-14159.

References

1. M. Pokrić, N. A. Thacker, M. L. J. Scott et al. "Multi-dimensional medical image segmentation with partial voluming." In *Proceedings of Medical Image Understanding and Analysis*, pp. 77–80. BMVA, 2001.
2. D. C. Williamson, N. A. Thacker, S. R. Williams et al. "Partial volume tissue segmentation using grey-level gradient." In *Proceedings of Medical Image Understanding and Analysis*, pp. 17–20. BMVA, 2002.
3. A. P. Dempster, N. M. Laird & D. B. Rubin. "Maximum likelihood from incomplete data via em algorithm." *Journal of the Royal Society* **39**, pp. 1–38, 1977.
4. M. Pokrić, N. A. Thacker & A. Jackson. "The importance of partial voluming in multi-dimensional medical image segmentation." In *Medical Image Computing and Computer-Assisted Intervention*, pp. 1293–1294. Springer-Verlag, 2001.

Consequences for imaging of the broadband nature of terahertz pulses

E Berry[a*], GC Walker[a], SW Smye[b], S Reed[a], JW Bowen[c], RE Miles[d], AG Davies[d]
[a]Academic Unit of Medical Physics and [d]School of Electronic and Electrical Engineering , University of Leeds, Leeds, UK. [b]Department of Medical Physics and Engineering, Leeds Teaching Hospitals NHS Trust, Leeds, UK. [c]Department of Cybernetics, School of Systems Engineering, The University of Reading, Reading, UK

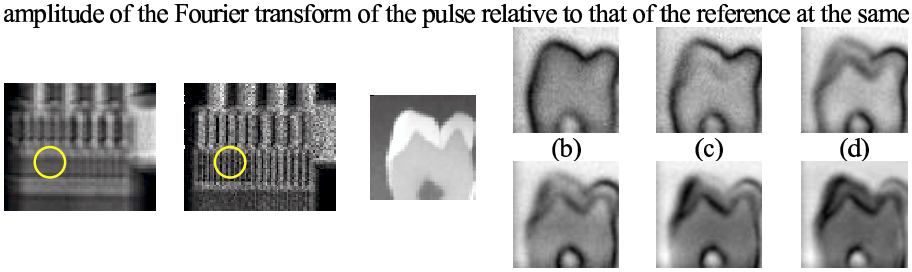
Abstract. In terahertz pulsed imaging, sets comprising images at the different frequencies in the terahertz pulse may be generated. These show the frequency-dependent effects of the subject on the radiation. In this pilot work, features of such image sets were studied with a view to their exploitation for image analysis. Data from test objects and from a slice of tooth were used. It was shown that, in a simplified case, edges could be tracked in a scale-space image generated from a terahertz image set, and it was concluded that edge focusing may be feasible. Frequency dependent artefacts seen at boundaries in transmission images were shown to be consistent with a geometrical optics description of within-beam interference. This new understanding provides the basis for future developments in physics-based terahertz image analysis.

1. Introduction

In terahertz pulsed imaging the basic measurement is a time series representing the amplitude of the pulse of terahertz radiation at a series of time points. The pulse peak has a FWHM of about 1 ps and the pulse is recorded for a period ~30 ps. The full acquisition consists of a reference pulse acquired without the subject, together with (for each pixel in the image) a pulse transmitted through, or reflected from, the subject. Terahertz pulsed images[1,2] are parametric and show the value of a parameter calculated from the time or frequency domain of the pulse data. In this work, the images used were calculated in the frequency domain and show the amplitude of the Fourier transform of the pulse relative to that of the reference at the same frequency.

(a) (b)

Figure 1. Transmission terahertz pulsed images of a spatial resolution test object at (a) 0.6 THz and (b) 2.8 THz



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

Figure 2. Images of 200 μm slice of tooth (a) X ray image, and transmission terahertz pulsed images at (b) 0.43 THz, (c) 0.53 THz, (d) 0.64 THz, (e) 0.75 THz, (f) 0.85 THz and (g) 0.96 THz.



Figure 3. Contrast detail test object at 1.6 THz, showing profile for scale-space analysis.

The pulse used in terahertz pulsed imaging comprises a range of frequencies, usually from about 0.5 THz to 3 THz. One consequence of this is that the width of the terahertz beam incident on the subject typically varies with frequency, which leads to a range of spatial resolutions in frequency domain images. In Figure 1, transmission images of a spatial resolution test object are shown, one calculated at 0.6 THz and the other at 2.8 THz. The same set of test bars is indicated by the circle in each case; they are resolved in the higher frequency image but not in the lower.

A second consequence of the existence of a range of frequencies is that the appearance of certain image artefacts is dependent on frequency. Figure 2(a) is an x-ray image of a 200 μm thick slice of tooth, this shows that the slice comprises regions of enamel (light grey) and dentine (mid grey). Frequency domain transmission terahertz images of the same slice of tooth in Figure 2 (b)-(g) demonstrate artefactually broad, dark boundaries between the enamel and dentine, and at their boundaries with air. These artefacts are more pronounced at some frequencies than at others, for example, compare the enamel/dentine boundary in Figures 2(b) and 2(g).

In this pilot work, the two effects were studied with a view to their exploitation for image analysis. Although the two effects are not independent, for the purposes of this investigation they have been treated entirely separately. The first study was prompted because sets of images where each member image has a different spatial resolution are reminiscent of the families of images generated for scale-space analysis [3,4]. Scale-space techniques are applied following the generation, from a single original image, of a family of derived images with successively coarser spatial resolutions, so similar techniques might also be applicable to the intrinsic families of images available from terahertz pulsed imaging. Little work has been published regarding boundary artefacts in terahertz pulsed images, so the second study was prompted by the possibility that a better

* e.berry@leeds.ac.uk

understanding of their origin would be useful for knowledge-driven image analysis. In the rest of this paper these two pilot studies are subtitled "scale-space" and "phase artefacts".

2. Background and theory

2.1 Scale-space

If edges are detected in a high resolution image the process is usually adversely affected by noise, but if a low resolution image is used, then the detected edges may not be located accurately. Bergholm[5] proposed the technique of edge focusing to overcome these difficulties. Edges are tracked from coarse to fine resolution, and so the process has high positional accuracy and is less affected by noise than processing at fine resolution. Fine to coarse resolution images are generated by blurring with a Gaussian operator, which has been shown[6] to have the correct mathematical characteristics for an operator to be used for edge focusing. In particular, the causality condition[7] is met, which means that no new detail is generated when the resolution is diminished. The set of images generated using the Gaussian operator at successively greater levels of blurring is known as a scale-space family. Whilst the families of images produced in terahertz pulsed imaging share the property of diminishing resolution, it is not known if they also satisfy the necessary condition of causality. The data in a scale-space family may be visualised with a scale-space image,[8] which is a plot of edge point location against spatial resolution. As resolution decreases, pairs of edge points merge to form closed contours confined to higher spatial resolution. If causality is satisfied, no new edges are plotted at low resolution that do not appear at higher resolution. Thus the scale-space image provides a preliminary visual assessment for determining the suitability of a family of images for edge focusing. New low resolution detail could occur in terahertz pulsed imaging if the frequency-dependent nature of the attenuation coefficient and refractive index of the materials led to contrast reversal, or to phase artefacts of the kind discussed in the second part of this paper, so exclusion of such confounding factors will be necessary for the pilot study. The aim was to determine (i) if edges could be tracked, and (ii) if spurious resolution was present, in a scale-space image calculated from a frequency domain set of terahertz pulsed images.

2.2 Phase artefacts

Normally, changes in phase arising purely from differing refractive index values, do not affect the measured amplitude of a transmitted signal. However it is possible to deliberately arrange for phase effects to be apparent in amplitude images. In microscopy, where it is common to image subjects that do not attenuate light but do affect its phase, techniques such as phase contrast microscopy have been developed. The phase difference between the subject and background are manipulated by optical techniques to be equivalent to $\lambda/2$, so the signals are completely out of phase and destructive interference results. In other imaging modalities, features arising from destructive interference due to optical path length differences have been exploited for measurement. In ultrasound imaging, interference within the beam is known as phase cancellation. The mechanism was first investigated by Rubin et al [9], and has been shown to result in a reduction in the measured signal in broadband ultrasonic attenuation measurements;[10] the effect of curvature of a bone surface has been successfully modelled.[11] In the case of the terahertz transmission images shown in Figure 2, the reduced amplitude at the boundaries could be explained by interference within the beam where the beam covers a boundary between two materials. This is a partial volume effect. In this situation, part of the beam travels through one material, and is delayed according to the refractive index of that material, whilst part of the beam is differently delayed in response to the refractive index of the second material. Minima are expected when the condition for destructive interference is satisfied, i.e. when the optical path difference is equivalent to an odd integer multiple of $\lambda/2$. Destructive interference at wavelength λ occurs if

$$\Delta(nd)_\lambda = (i+\tfrac{1}{2})\lambda \qquad \text{(Equation 1)}$$

where i is integer ≥ 0, and $\Delta(nd)_\lambda$ is the optical path difference (Figure 4) arising from differences in refractive index (n), physical path length (d) or both.

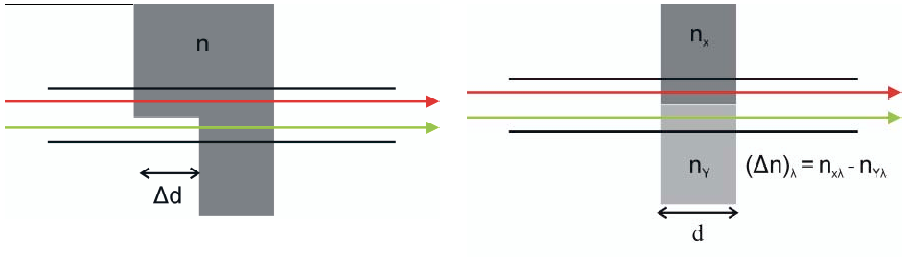


Figure 4. Alternative geometries for within-beam optical path length differences. The red and green arrows indicate different paths within the beam that result in within-beam interference. (a) Physical path length difference of Δd in sample of uniform refractive index n. (b) Refractive index difference Δn in sample of constant thickness d. d and Δd are constant with frequency, but n may be frequency-dependent.

If n is constant with frequency, for a physical step in a single material, Equation 1 reduces to $\Delta(d)_\lambda = (i+½)\lambda/n$. From this expression, the first two minima for a given step are predicted to occur at λ_0 and λ_1, where

$\lambda_0 = 3\lambda_1$ (Equation 2) i.e. $f_1 = 3f_0$ (Equation 2a)

Furthermore, if two steps of depths Δd_A and Δd_B are present, it can be shown that if F_{sep} is the separation between consecutive minima for a given step size, then for constant n

$\Delta d_A/\Delta d_B = F_{sep\,B}/F_{sep\,A}$ (Equation 3)

From previous measurements in the terahertz region we have shown[12] that refractive indices are approximately constant with frequency above 0.5 THz for some materials such as TPX and Duraform (Nylon 12), but it would not be a valid assumption to apply to tissues.

For a boundary between two materials with different refractive indices and constant thickness d, Equation 1 reduces to $\Delta(n)_\lambda = (i+½)\lambda/d$. From this expression, the first two minima for a given refractive index boundary are again predicted to occur at λ_0 and λ_1, where $\lambda_0 = 3\lambda_1$. In this case, it is assumed that the refractive index *difference* between the two materials, $\Delta(n)_\lambda$, is that same at each minimum. For a first approximation, this is a reasonable assumption[12] for both tissues and man-made materials. The aim of this study was to confirm the hypothesis that the boundary artefacts satisfy Equations 2 and 3 and that they are thus consistent with interference within a beam that has covered both sides of a boundary between materials.

3. Methods

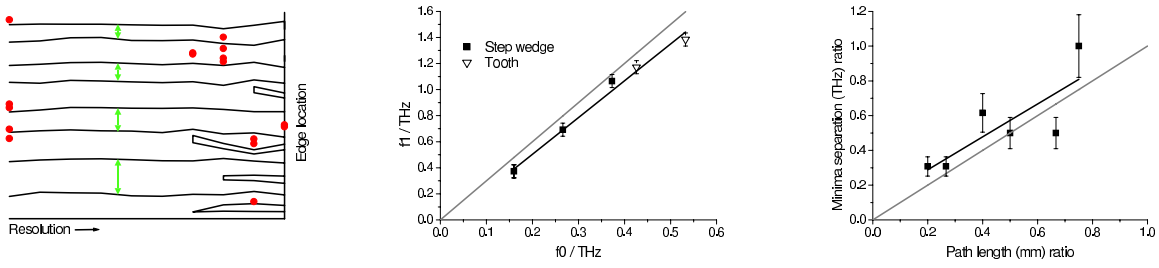
3.1 Scale-space

Transmission mode terahertz pulsed imaging data were acquired, and image sets generated, from a test object with circular contrast-details (Figure 3). The substrate was TPX, with gold contrast-details. This was a choice that minimised the likelihood of spurious resolution arising from frequency-dependent attenuation characteristics, as the absorption coefficient of TPX is negligible and gold reflected, rather than absorbed, the radiation. Similarly, phase artefacts were avoided because the attenuation by the gold was by reflection, so TPX was the only material through which the radiation was propagated, even at boundaries. At each pixel 32 points separated by 0.08 ps were recorded, corresponding to a frequency resolution of 390 GHz after Fourier transformation. Scale-space images were generated from a profile (Figure 3) through the image set from 0.4 THz to 3.9 THz. Edges were defined as zero-crossings in the second derivative of the profile, and those clearly seen in the mid-frequency images were manually tracked through the increasing resolution region of scale-space. Edges violating the causality condition were identified.

3.2 Phase artefacts

Transmission mode terahertz pulsed imaging data were acquired, and image sets generated, from a step wedge of Duraform,[13] and from a 200 μm thick slice of tooth held in air. For the step wedge, at each pixel, 128 points were acquired at 0.15 ps intervals, giving a frequency resolution of 52 GHz. The tooth slice image was 56 by 56, and each time series comprised 64 points separated by 0.15 ps, giving a frequency resolution of 104 GHz. The step wedge had path length differences of 0.4 mm, 1 mm, 1.5 mm and 2 mm. Measurements of pixel value were made at each frequency from profiles of the steps, and through boundaries between air, enamel and dentine in the tooth images. Minima were identified to obtain f_0 and f_1, f_0 was plotted against f_1 and the slope compared the value of 3 predicted in Equation 2a. For the step wedge, the validity of Equation 3 was tested using the known step sizes. Errors were calculated from the uncertainty in the localisation of the frequency minima.

4. Results



(a) (b) (c)

Figure 5. (a) Scale-space image from the profile in Figure 3. The green arrows indicate where the circular details appear. The red points indicate detected edges that violate the causality condition. (b) Frequencies for the first two destructive interference minima, black line is fitted to measurements and grey line shows Equation 2a. (c) Experimental results for step wedge, black line is fitted to measurements and grey line shows Equation 3.

4.1 Scale-space

The scale-space image, Figure 5(a), had edges that were continuous and could be tracked to high resolution. However, even for this simple object, points violating the causality condition were found.

4.2 Phase artefacts

The expected cyclical variation in pixel value with frequency was observed for boundaries in both step wedge and tooth, except for the dentine/enamel border where f_1 was beyond the range of frequencies acquired. The slope of the experimental line in Figure 5(b) was 2.82 ± 0.15. Results from the step wedge, shown in Figure 5(c), were consistent with the prediction of Equation 3, and the slope of the experimental line was 0.95 ± 0.35.

5. Discussion

5.1 Scale-space

Edge focusing in terahertz pulsed imaging families may be possible, especially where the frequency dependence of the material properties is known. The most suitable resolution at which to start the focusing process requires further investigation. The example chosen was considerably over-simplified as the test object did not mimic the frequency-dependent optical properties of tissue, which could result in explainable spurious resolution. More complete modelling is indicated, together with extension to reflection imaging, which is likely to be the modality used clinically.

5.2 Phase artefacts

The results in Figures 5(b) and 5(c) show that the simple geometrical optics model for partial volume effects explains a major contribution to the artefacts seen at boundaries in terahertz pulsed images. A small contribution may arise from deflection out of the beam,[14] but although investigation has begun[15] scattering in terahertz imaging remains poorly understood. This pilot work used only two data sets, so there is scope for a more complete study. Inclusion of the expected frequencies of constructive interference may be valuable and work is needed to extend the argument to reflection imaging. The effects will be exploitable for image analysis by, for example, the application of calibration boundaries, incorporation of variation with frequency into cluster analysis, or the use of models of frequency-dependent edge profiles for sub-pixel localisation. In-beam interference from reflections at the edge of a step[16] has previously been exploited in an ultrasonic rangefinder. An analogous instrument could be built using a tuneable continuous wave terahertz source. Although recent work[17] has shown that additional frequency-dependent effects are present in terahertz pulsed imaging systems, these effects were not considered in this analysis. These arise from the use of a convergent beam to interact with the sample and also because the detection efficiency is greater at the higher frequencies in the pulse.

Acknowledgements

We are grateful for financial support from EPSRC (GR/N39678, GR/P00086/01) and EU (Teravision - IST-1999-10154), and acknowledge the contributions of AJ Fitzgerald, T Loeffler, K Siebert, DJ Wood, S Strafford, NN Zinov'ev, JM Chamberlain, MA Smith.

References

1. B.B.Hu and M.C.Nuss "Imaging with Terahertz Waves", *Optics Letters* 20, pp. 1716-1719, 1995.
2. D.M.Mittleman, R.H.Jacobsen, and M.C.Nuss "T-ray imaging", *IEEE J Sel Top Quant Electron* 2, pp. 679-692, 1996.
3. A.P.Witkin "Scale-space filtering", *Proc Int Joint Conf Artificial Intell (Karlsruhe, Germany)*, pp. 1019-1023, 1983.
4. B.M.ter Haar Romeny. *Geometry-Driven Diffusion in Computer Vision.* Kluwer Academic Publishers, Dordrecht/Boston/London , 1994.
5. F.Bergholm "Edge Focusing", *IEEE Trans Pattern Anal Machine Intell* 9, pp. 726-741, 1987.
6. J.J.Koenderink "The structure of images", *Biological Cybernetics* 50, pp. 363-370, 1984.
7. P.Perona and J.Malik "Scale-space and edge detection using anisotropic diffusion", *IEEE Trans Pattern Anal Machine Intell* 12, pp. 629-639, 1990.
8. M.Sonka, V.Hlavac, and R.Boyle *Image processing, analysis and machine vision.* Brooks/Cole Publishing Company, Pacific Grove , 1999.
9. J.M.Rubin, R.S.Adler, J.B.Fowlkes et. al. "Phase Cancellation - A Cause of Acoustical Shadowing at the Edges of Curved Surfaces in B-Mode Ultrasound Images", *Ultrasound Med Biol* 17, pp. 85-95, 1991.
10. G.W.Petley, P.A.Robins, and J.D.Aindow "Broad-Band Ultrasonic-Attenuation - Are Current Measurement Techniques Inherently Inaccurate", *Br J Radiol* 68, pp. 1212-1214, 1995.
11. C.M.Langton and M.Subhan "Computer and experimental simulation of a cortical end-plate phase cancellation artefact in the measurement of BUA at the calcaneus", *Physiol Meas* 22, pp. 581-587, 2001.
12. G.C.Walker "Modelling the propagation of terahertz radiation in biological tissue", PhD Thesis: The University of Leeds, 2004.
13. A.J.Fitzgerald, E.Berry, R.E.Miles et. al. "Evaluation of image quality in terahertz pulsed imaging using test objects", *Phys Med Biol* 47, pp. 3865-3873, 2002.
14. T.Loffler, T.Bauer, K.J.Siebert et. al. "Terahertz dark-field imaging of biomedical tissue", *Optics Express* 9, pp. 616-621, 2001.
15. D.Mittleman, J.Pearce, and Z.Jian "Imaging with multiply scattered terahertz pulses", *Med Phys* 30, pp. 1540-1540, 2003.
16. G.C.Bakos and N.F.Tsagas "Ultrasonic Rangefinder - Edge-Detection and Enhancement Method", *Int J Electron* 77, pp. 511-516, 1994.
17. J.W.Bowen, G.C.Walker, S.Hadjiloucas et. al. "The consequences of diffractively spreading beams in ultrafast THz spectroscopy" In *Proc Joint 29th Int Conf Infrared Millimeter Waves and 12th Int Conf on Terahertz Electron (Karlsruhe, Germany)*, 2004.

Elasticity Reconstruction from Displacement and Confidence Measures of Ultrasound Image Sequence

Junbo Li, J. Alison Noble*
Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ

Abstract: Ultrasonic elasticity imaging (UEI) is showing promise as a new way of imaging the elasticity of soft tissues to provide additional diagnostic information to B-mode ultrasound to help distinguish healthy from diseased tissue. The premise of the technique is that elastic characteristics of soft tissue may differ in various pathological states. We propose a multi-frame finite-element based scheme to estimate the elasticity of soft tissue from an ultrasound image sequence. Although multi-frame processing is computationally intensive, the presented method shows promise to give significantly higher accuracy for elasticity imaging than previously reported work by ourselves [1] and other published methods.

1. Introduction: Ultrasound elasticity imaging techniques provide new opportunities for detection and diagnosis of cancers in the breast, prostate, coronary artery and other sites by assessing the elastic characteristics of soft tissue. Most of these techniques can be classified into the direct or inverse approaches [2-7]. The inversion elasticity imaging approaches have been widely investigated because of their robustness, while the noises of displacement estimate in solving the ill-posed inverse problem has become an impediment to restrict its diagnostic effectiveness, especially for ultrasound images with limited spatial and contrast resolution [4-7].

In this paper, the potential benefits of elastography from ultrasound video sequences have been explored. A 3-D freehand ultrasound imaging system is used to deform tissue with certain contact forces at the same time as imaging. The Displacement distributions over consecutive B-mode images are calculated using maximum likelihood-based block matching method [8]. Confidence measures of displacement estimation are employed as weighting factors to evaluating the objective functions of the inverse problem. Averaged strain image of the acquired sequence is used as *a prior* knowledge of the tissue relative stiffness distribution to constrain the inverse solutions. A finite-element based inverse split-and-merge strategy [1] is presented for elasticity reconstruction. The L_2-error between the target and reconstructed Young's modulus is found to be less than 1% for a single inclusion model even with 5% noise level. Experimental results on a gelatin phantom are presented. Although the results are based on B-mode image sequence, improvements are expected by using RF data.

2. Method

A. Data Acquisition: A 3-D freehand Ultrasound system was developed. It consists of an ultrasound imaging machine (Sonos 5500, Agilent Technologies, USA), optical tracker (Polaris Hybrid, Northern Digital Inc, Canada), composite probe (defined below), frame grabber (Meteor II-MC, Matrox Imaging, Canada) and host PC. The overview of the system is shown in figure 1. A 6-axis force transducer (Mini 40, ATI Industrial Automation, USA) and 4 infrared LED were mounted on a 7.5 MHz ultrasound linear array transducer (Hewlett Packard L7540) and integrated together in an enclosing box as the composite probe (figure 2).



Figure 1. Freehand ultrasound system.

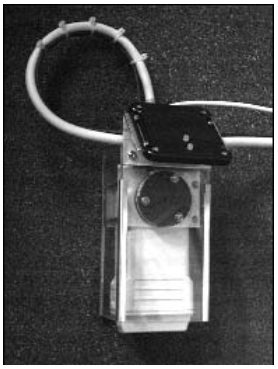


Figure 2. Composite probe

Before ultrasonic scanning, a spatial calibration procedure was performed to enable the actual position and orientation of the B-scan images with respect to the world coordinate system to be determined [9]. Ultrasonic coupling gel (Parker Laboratories, NJ, USA) was used to couple the transducer and the specimen. At the beginning of the scan, the transducer was placed perpendicularly to the surface of the specimen with minimal

* Junbo Li and J. Alison Noble are with the Medical Vision Lab, Robotics Group, Department of Engineering Science, University of Oxford. Email: {larry, noble}@robots.ox.ac.uk, www: http://www.robots.ox.ac.uk/~mvl/external/mvl-html/

force and then pushed down and released slowly to compress the specimen. B-scan images, positions of the transducer and contact forces were acquired simultaneously during the scan.

B. Maximum Likelihood Displacement Estimation: Displacements over consecutive B-mode images are calculated by a maximum likelihood-based block matching method [8]. The best match provides the most likely observed displacement between the two images. The cross-correlation coefficients of displacement are calculated as a measure of confidence of the estimate.

Let $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ be matched blocks from two consecutive frames $\boldsymbol{x}$ and $\boldsymbol{y}$ and the displacement vector denoted by $\boldsymbol{v}_i(u,v)$. The maximum likelihood (ML) estimation of $\boldsymbol{v}_i$ is given by

$$v_i^{ML} = arg \max_{v_i} p(x_i | y_i, v_i) \qquad (1)$$

where indices $\boldsymbol{i}$ is an index over all possible blocks. .In ultrasound, when the speckle is fully developed, the noise is multiplicative and follows a Raleigh pdf. We can assume blocks $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ are corrupted by independent multiplicative Raleigh distributed noise, representing uniform dense speckle. By taking the natural logarithm, the multiplicative noise is transformed to additive. Maximizing the pdf is equivalent to maximizing the following CD_2 objective function [8]

$$E_i^{CD_2}(v_i) = \sum_j \left\{ \tilde{x}_{ij} - \tilde{y}_{ij} - \ln(\exp(2(\tilde{x}_{ij} - \tilde{y}_{ij})) + 1) \right\} \qquad (2)$$

where indices $\boldsymbol{j}$ is an index of pixels in the block.

This similarity measure assumes that the noise distribution in both of the blocks is the same. An intensity normalization must be conducted before calculation of CD_2, where the original intensity values of $\tilde{x}_i$ and $\tilde{y}_i$ are transformed into new values by subtracting the mean and dividing the standard deviation to make sure the two blocks have the same mean and variance. This modified similarity measure is denoted as CD_{2bis} and is used in Singh's block-matching framework [10] for the estimation of the displacement field.

Following [8], the probability mass function is defined as

$$R_c(u,v) = \frac{1}{Z} \exp\left\{ k \frac{E_c(u,v) - m}{(2n+1)^2} \right\}, \qquad -N \le u, v \le N \qquad (3)$$

where $E_c(u,v)$ is the CD_{2bis} similarity measure of a $2n+1$ long block $\boldsymbol{W}_c$ for each candidate displacement (u,v) and m is its maximum in the search window $\boldsymbol{W}_s$. k is a parameter of constant value to adjust the maximum response in $\boldsymbol{W}_s$ close to unity. The velocity estimate is defined as the mean of a thresholded version of R_c

$$u_c^h = \frac{\sum_u \sum_v R_c^h(u,v) u}{\sum_u \sum_v R_c^h(u,v)} \qquad v_c^h = \frac{\sum_u \sum_v R_c^h(u,v) v}{\sum_u \sum_v R_c^h(u,v)} \qquad (4)$$

where

$$R_c^h(u,v) = \begin{cases} R_c(u,v) & \text{if } R_c(u,v) \ge \alpha \\ 0 & \text{otherwise} \end{cases}, \qquad \alpha = \hat{m} - h(\hat{m} - \check{m}) \quad \text{with} \quad h \in [0,1]$$

$\hat{m}$ and $\check{m}$ are the maximum and the minimum of the probability mass function R_c respectively.

C. Forward Problem in Elasticity Imaging: A tissue elasticity forward model is built to calculate the theoretical displacements. We assume that the transducer slowly compresses the tissue so that tissue deformation is regarded as quasi-static. We also assume that the elasticity distributions remain constant during compression. This will hold true in the elastic-plastic case. The governing equations for such a situation have been described in many papers [1, 5-7, 11]. Given boundary constraints and a finite element mesh, a finite element model is built and the theoretical displacements throughout the body can be calculated using finite-element analysis. In this study the measured contact force is assumed to be uniformly distributed along the contact surface and applied to the finite element model that represents the mechanical behavior of the tissue. The surface displacements of the tissue and some constraints on the movement of the surfaces are employed as the boundary conditions of the finite element model.

D. Inverse Problem Using Multiple Images: Conventional inverse problem of UEI can be stated as to find the smallest distance between measured displacements from 2 images and their counterparts described by the forward finite element model, in which the elasticity parameters act as the state variable, the displacements act as system output variable and is fully controlled by the model physics and the initial and boundary conditions. We modified this standard approach in [1] by using an image sequence and a split-and-merge strategy. Currently we make a further improvement by including a measure of confidence in the reconstruction process. Thus,

$$\hat{E} = \arg\min\{F(E)\}, \qquad F(E) = \sum_{i=1}^{M} F_i(E) = \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{N_i} W_{ij}\left\|u_{ij}(E) - \hat{u}_{ij}\right\|^2 \tag{5}$$

where $\hat{E}$ is the estimated elasticity distribution, $F(E)$ is the objective function, M is the number of image-pairs for displacement calculation, N_i is the number of observation points of the ith image pair, $u(E)$ is the theoretical displacement vectors for a given finite element model, and $\hat{u}$ is the observed displacement vector from subsequent images. The term W_{ij} is the weighting factor that controls the contribution of observation (i, j) to the reconstruction problem and is related to the confidence of displacement estimates. The idea is that the "trusted" data contributes more to the objective function than the less confident estimated data.

The inverse problem is then recast as a non-linear optimization problem. A modified Levenberg-Marquardt algorithm and an active set strategy are used to solve the nonlinear least square problems subject to upper and lower bounds on Young's modulus [1, 11-12]. The non-linear optimization solution is achieved using a Fortran routine DBCLSJ in IMSL libraries. The upper and lower bounds on Young's modulus are set according to the elastic range of specimen and the Jacobian matrix is calculated by its finite difference approximation.

E. Split-and-Merge Strategy: A split-and-merge strategy is employed over the image sequence in which strain images are used to provide *a priori* knowledge of the relative stiffness distribution of the tissue to constrain the inverse problem solution. The elasticity reconstruction problem is then solved iteratively by a coarse-to-fine approach as in [1]. First, the strain distributions are computed from differences of the observed displacement fields of the whole sequence. An averaged strain distribution is then obtained. A pyramid of regions of constant strain is then estimated from the strain image using coarse-to-fine fuzzy K-means image classification. This classification will contain true regions and falsely classified regions due to strain artifacts. Assuming each labeled region has the same value of Young's modulus (ie assuing uniform stress), the inverse problem is solved to get the elasticity distribution of the tissue. The theoretical displacement of the entire image is then calculated by solving the forward problem using the estimated elasticity distribution. If the normalized differences of the reconstructed elasticity in different regions is small, these regions merge into one region. If the mean squared error (MSE) between the measured displacements and their theoretical value is bigger than a pre-set tolerance, we increase the class number (ie., the number of region). The hierarchical process continues with a new set of labeled data from the pyramid with increased class number. The iterations stop when the error between the measured and calculated displacements is smaller than a pre-set tolerance.

3. Results: The new method was firstly studied by a simulated single inclusion model. Uniform pressure loads were applied on the upper surface of the model. Axial displacements on the lower surface were restricted to zero. The center point on the lower surface was regarded as a "steady point" and its displacement vector set to zero. The Young's modulus for the inclusion and the background is 0.02 MPa and 0.2 MPa respectively.

The reconstructed elasticity distributions with different levels of noises and image pairs are shown in Table 1. We observe that the L_2-errors reduce significantly with an increase in the number of image pairs. Note also that the elasticity distributions can be recovered even with poor SNR. Compared with previous work [7], in which the reconstructed elasticity of a single inclusion model is underestimated by about 40% with 3% noise on the displacement estimation, our method is superior, with a lower L_2-error.

Table I. Elasticity Reconstruction under Different Noise Level and Image Pairs

Number of Image pairs	L_2-errors of the reconstructed Young's modulus with different noise levels			
	0% Noise	5% Noise	10% Noise	20% Noise
1	0	3.25%	20.03%	84.84%
5	0	2.02%	13.91%	31.26%
10	0	1.26%	10.70%	17.12%

A gelatin block-shaped phantom was built from gelatin (Sigma G2500, 300 Bloom, Type A) and distilled water. A cylindrical inclusion was constructed 1 mm deep from the upper surface and in the lateral center of the phantom, with a diameter of 33 mm. The Young's modulus of the background material should be 16 times that of the inclusion. Talcum powder was added to the surrounding gelatin of the inclusion as scatterers to increase the contrast between the inclusion and its background in the B-scan images.

The relationship between estimated relative elasticity and applied contact force is shown in Fig.3 (a). The relative elasticity is defined as the ratio of the reconstructed Young's modulus of the background to that of the soft inclusion. It decreases with an increase in the contact force. This may be due to inaccurate estimation of the displacement resulting from the small deformation of the phantom with a small contact force. The estimated

relative elasticity using the image pair that undergo the largest deformation is close to the theoretical value of 16. Figure 3 (b) shows the estimated relative elasticity from an image sequence acquired during phantom deformation. The first and last images in the sequence are related to the minimum and maximum contact forces. Image pairs are carefully selected from the sequence with nearly equal contact force increments. No significant trend is observed from the results to support the hypothesis that the estimated elasticity of the phantom is more accurate with more image pairs. This is probably because the mechanical characteristics of the phantom are well-modeled and its structure can be clearly observed from the B-scan images. This greatly reduces the noise in the theoretical and observed displacement distribution and makes the reconstructed relative elasticity stable. It can be also seen from Fig.3 (b) that the calculated relative elasticity is larger than its theoretical value when 10 and 20 image pairs are used. This is due to the poor SNR of displacement estimation when there is a small displacement between sequential images. This highlights that it is important to carefully select the image pairs to ensure good precision of displacement calculation.

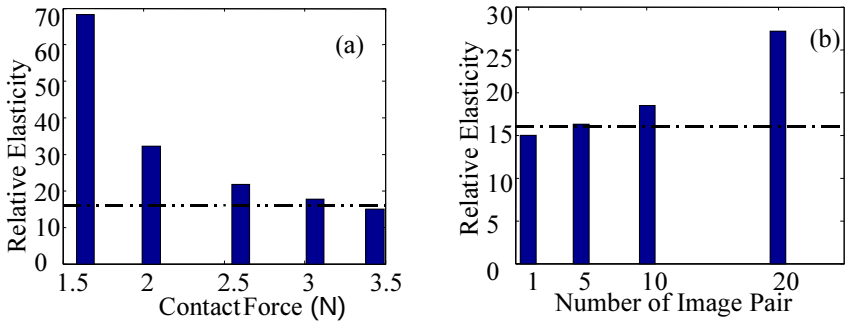


Figure 3. Estimated Relative Elasticity of the cylinder phantom with different (a) Contact forces, and (b) Number of Image Pairs. The dash-dot line indicates the theoretical value of relative elasticity.

4. Conclusion: We have introduced a new method for elasticity reconstruction from a sequence of ultrasound images. The inverse elasticity problem is constrained by the strain image to render the solution to be stable. Multiple image pairs under different loads are used to get good reconstruction results even with poor SNR images. Simulations and experiments on a gelatin phantom have demonstrated the feasibility of the approach. Future work will seek to apply the method to clinical data and adapt the approach to work with RF data rather than B-mode images.

References

[1] J. Li, J. A. Noble and M. Burcher, "Inversion Elasticity Reconstruction of Soft Tissue Using Split-and-Merge Strategy from Strain Map of Ultrasound Image Sequence", Proc. of the IEEE International Ultrasonics Symposium, Hawaii, USA, pp. 1927-1930, 2003.

[2] J. Ophir, I. Cespedes, H. Ponnekanti, Y. Yazdi, and X. Li, "Elastography: a quantitative method for imaging the elasticity of biological tissues," *Ultrason. Imag.*, vol. 13, pp. 111-134, 1991.

[3] M. O'Donnell, A. R. Skovoroda, B. M. Shapo and S. Y. Emelianov, "Internal displacement and strain imaging using ultrasounic speckle tracking," *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, vol. 41, pp. 314-325, 1994.

[4] K. R. Raghavan, and A. E. Yagle, "Forward and inverse problems in elasticity imaging of soft tissue," IEEE Trans. Nucl. Sci., vol.41, no.4, pp.1639-1647, 1994.

[5] F. Kallel and M. Bertrand, "Tissue elasticity reconstruction using linear perturbation method", IEEE Trans. Medical Imaging, vol 15, no. 3, pp. 299-313, 1996.

[6] E. E. Van Houten, K. D. Paulsen, M. I. Miga, F. E. Kennedy and J. B. Weaver, "An overlapping subzone technique for MR-based elastic property reconstruction", Magnetic Resonance Medicine, 42: 779-786, 1999.

[7] A. A. Oberai, N. H. Gokhale and G. R. Feijóo, "Solution of inverse elasticity problems using the adjoint method: the incompressible case", Technique Report 02-009, Boston University, 2002.

[8] D. Boukerroui, J. Alison Noble and M. Brady, "Velocity Estimation in Ultrasound Images: A Block Matching Approach", Proc. of the 18th International Conference of Information Processing in Medical Imaging, IPMI 2003, Ambleside, UK, 586-598, 2003.

[9] G. Xiao, "3-D freehand ultrasound imaging and image analysis of the breast", Ph.D. thesis, Oxford university, 2001.

[10] A. Singh and P. Allen, "Image-Flow Computation: An estimation - Theoretic Framework and Unified Perspective", CVGIP: IU, vol. 56(2), pp 152-177, 1992.

[11] L. Han, J. A. Noble and M. Burcher, "The Elastic Reconstruction of Soft Tissue", IEEE International Symposium on Biomedical Imaging: Macro to Nano, 2002.

[12] IMSL Math/ Library Manual, Visual Numerics Inc, 1997.

Correction of misaligned slices in multi-slice MR cardiac examinations by using slice-to-volume registration

A. G. Chandler[a], T. Netsch[b], J. A. Schnabel[a], R. Razavi[a], D.L.G. Hill[a], D. J. Hawkes[a]
[a]Imaging Sciences Group, Guy's Hospital, King's College London, London, UK
[b]Philips Research laboratories, Roentgenstrasse 24, Hamburg, Germany

Abstract. One of the main challenges with Magnetic Resonance (MR) cardiac image acquisition is to account for cardiac motion due to respiration. A popular technique to reduce respiratory motion is to perform short axis (SA) MR multi-slice acquisitions in which a patient holds his/her breath multiple times during the scan. This paper explores the feasibility of using normalized mutual information based rigid slice-to-volume registration to correct for misalignments of slice stacks in such images due to differing breath hold positions.

1 Introduction

The slice-to-volume registration application that we address in this paper is the accurate and robust rigid registration of MR cardiac slice stacks obtained from a multi-slice short-axis (SA) acquisition to an image obtained from a 3D volume acquisition. It is not possible to acquire sufficient slices to cover the entire heart in a single breath hold, even by using the latest improvements in acquisition technology such as Steady State Free Precession (SSFP) imaging and sensitivity encoding. These acquisitions typically involve three or more different breath holds. Inconsistencies in the breath hold position can cause misalignment of slices relative to each other. Slice-to-volume registration can correct this misalignment and will produce anatomically more correct images, which, for example, will provide more accurate ventricular function analysis.

2 Purpose

Medical imaging has an important role in the management of patients with heart disease. Cardiac MR is currently the gold standard for the assessment of ventricular function [1]. One of the main challenges with cardiac image acquisition is to account for cardiac motion due to respiration. If this motion is not corrected, it can lead to severe artifacts in the final images. Numerous techniques for motion compensation have been proposed for use with MR acquisitions, including respiratory gating [2] and model based methods [3,4], but poor reproducibility of the respiratory cycle and the complex motion of the heart with respiration [5] means that breath holding is still the most reliable and widely used approach.

In view of the spatial and temporal resolution needed for the standard SSFP imaging of the heart, it is only possible to acquire between 1 and 3 slices in a single breath hold. The most common and important sequence used in cardiac MR are short-axis SSFP stack images, which usually consists of 10-14 parallel slices. This sequence is used to look at global and regional ventricular function. During the acquisition of the SA stack image, the subject is asked to hold their breath multiple times at the same breath hold position. At each breath hold, only a proportion of the image slices are acquired. The major draw back with such a method is that the quality of the multi-slice image volume obtained depends greatly on how consistently a subject can hold his or her breath at the same position. If the position of the heart in the different breath holds is very different, there will be errors introduced when the images are analyzed. Although fast undersampling acquisition techniques [6] have been developed to try to obtain the similar image resolution in a single short breath hold, they have not yet been shown to provide equivalent information, and are not in clinical use.

Recently it has been possible to acquire three-dimensional (3D) SSFP volumes of the heart for 3-7 cardiac phases using a multi-chunk sequence for delineation of cardiac anatomy especially in patients with congenital heart disease. This acquisition is performed in a straight axial plane which allows us to obtain a 3D volume of the heart with good spatial resolution. Using these 3D volumes, we believe that retrospective slice-to-volume registration is capable of correcting for the misalignments caused by differing breath hold positions in the SA stack image which will result in a anatomically more correct image for better analysis, including wall motion and left ventricular volume analysis. This is the first time, to our knowledge, that slice-to-volume registration has been used on human subjects to correct for inconsistencies in the position of breath holds.

3 Materials

We acquired 3 data sets for the experiments in this paper. The first was a high-resolution 3D SSFP MR axial cardiac volume (256x256 pixels, 120 slices, 1.25mmx1.25mm in-plane sampling with 1mm slice thickness, 4 phases, acquired on a Philips Gyroscan Intera 1.5T MR system). The two other data sets were MR multi-slice SA images. (256x256 pixels, 12 slices, 1.25mmx1.25mm in-plane sampling with 8 mm slice thickness, 25 phases, acquired on a Philips Gyroscan Intera 1.5T MR system). Of the 12 slices, all pairs of consecutive slices were acquired at different breath holds. One of the two image volumes was motion-free, obtained by asking a cooperative volunteer to hold their breath at approximately the same exhale position for each of the 6 breath holds. This produced 6 stacks (containing two slices per stack) that had good anatomical alignment with respect to each other. The other image volume was a truly misaligned image in which the individual acquisitions are misaligned by asking the volunteer to hold their breath at relatively different exhale positions for each of the 6 breath holds. This image is similar to the image that would be acquired of a patient who has difficulty holding their breath at the same position.

4 Methods and experiments

The accuracy and robustness of rigid slice-to-volume registration was evaluated in two sets of experiments. The aim for both experiments was to correct for misaligned cardiac anatomy in multi-slice SA images by registering stacks of two slices to a high-resolution 3D MR axial cardiac volume using normalised mutual information (NMI) at the end diastolic phase. The experimental procedures carried out will be described for the two sets of experiments.

4.1 Simulation experiment

For this experiment, we simulated a misaligned SA image and tried to recover the misalignment. To do this, we used the motion-free MR multi-slice SA image. The stack positions in this image were used as the gold standard stack position. Regions within the SA image which were not of interest, such as the chest wall and spine were excluded using a coarse manual segmentation. We then artificially misregistered the 6 stacks from the end-diastolic phase image by random amounts to allow us to simulate a subject holding their breath at relatively different positions for each breath hold. To simulate realistic misaligned breath hold positions, the results from a study of the motion of the heart due to respiration [5] were used to limit the misregistration applied. For each of the 6 stacks, 20 random misregistrations constrained by these limits were applied. The simulated misaligned slice stacks were then rigidly registered, using NMI, to the end-diastolic phase image of the high-resolution 3D MR axial cardiac volume to try to recover the applied misregistrations.

By using the end-diastolic phase of the multi-slice SA image and the 3D volume image, the differences in cardiac anatomy due to heart deformation over the cardiac cycle could be ignored allowing us to correct for just the change in position of the heart due to differing breath hold positions using a rigid body transformation.

4.2 Correction of truly misaligned SA images

For this experiment, the 6 stacks from the end-diastolic phase of the misaligned SA image were segmented for the reasons explained above in section 4.1, and rigidly registered separately to the end-diastolic phase image of the 3D volume in order to correct the misalignment of the stacks. Once the rigid body transformations were obtained for each end-diastolic stack, they were use to realign the corresponding misaligned stack in the other 24 phases of the multi-slice SA image.

5 Results

This section presents the evaluation of the accuracy and robustness of slice-to-volume registration for the correction of misaligned SA stacks. A study on how the analysis of the left ventricular volume is affected due to misalignment of the stacks and the improvements made with correction by registration is also included.

5.1 Simulation experiment

Visual inspection of the motion-free multi-slice acquisition indicated that the 6 different stacks were well aligned. To validate the accuracy and robustness of the simulated experiment, the *mean imaged patient voxel displace-*

ment [7] (MIPVD) was used to determine how well the registration recovered the 20 misregistrations for each of the 6 stacks. The MIPVD is the mean of the Euclidian distances of each voxel centre position transformed by the registration transformation, and a gold standard transformation. The 6 stack positions of the motion-free MR multi-slice SA image were used as a gold standard. The mean, minimum, and maximum MIPVD for each of the 6 stacks before and after registration are shown in table 1.

	MIPVD in mm					
	Before registration			**After registration**		
Stack number	**Mean**	**Minimum**	**Maximum**	**Mean**	**Minimum**	**Maximum**
1	6.99	1.83	11.45	2.21	1.68	2.41
2	6.99	1.85	11.40	2.10	1.63	3.29
3	7.01	1.82	11.52	1.58	1.27	2.05
4	7.01	1.83	11.49	1.26	0.82	2.02
5	7.03	1.81	11.59	1.11	0.88	1.46
6	7.03	1.80	11.65	1.12	0.89	1.28

Table 1: The mean, minimum, and maximum MIPVD before and after registration recovered the 20 artificial misregistrations, for each of the 6 stacks of the motion-free multi-slice SA end-diastolic image.

The results in table 1 indicate that slice-to-volume registration is sufficiently robust and accurate within sub-voxel accuracy to be able to recover the expected misalignments of the slice stack position in a realistic situation.

5.2 Correction of truly misaligned SA images

To determine how accurately slice-to-volume registration has corrected the misalignment of the 6 stacks in the misaligned SA image, we visually compared the alignment of the left ventricular wall before and after registration to the left ventricular anatomy obtained for the motion-free SA image. Figure 1 shows a visualization of these segmentations for cardiac phase 14.

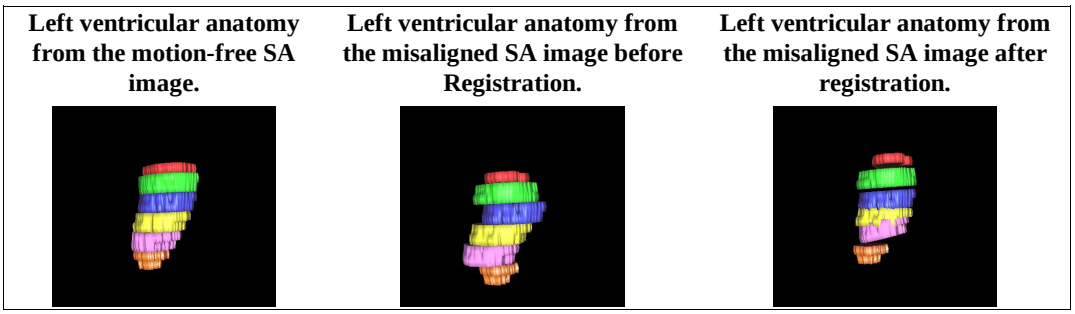


Figure 1: Visualization of the left ventricular wall from the motion-free SA image, the misaligned SA image before and after registration, respectively, for cardiac phase 14. The different colours/grey values represent the 6 stacks acquired at different breath holds.

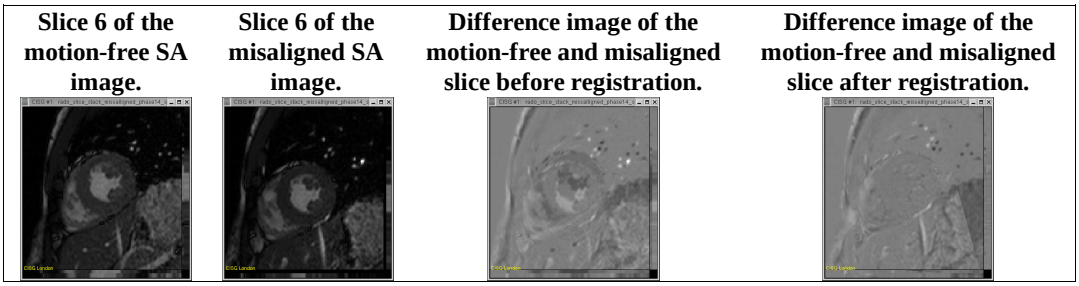


Figure 2: Columns 1 and 2 show motion-free and misaligned short axis slices (slice 6) of the heart for cardiac phase 14. Column 3 shows the difference image between the motion-free and misaligned SA slice before registration, and column 4 shows the difference image between the motion-free and misaligned slice after registration.

From figure 1, it can be seen that the wall of the left ventricle (LV) in the motion-free SA image is a good representation of the ventricular anatomy whereas the wall of the LV of the misaligned SA image before registration shows significant displacement due to misregistration. The wall of the LV of the misaligned SA image after regis-

tration depicts a smoother ventricular surface, very much like the left ventricular surface obtained from the motion-free image. To confirm how well the registration performed, slices of the motion-free SA image were compared, by means of difference images, to the corresponding slice positions of the misaligned SA image before and after registration, for all 25 cardiac phases (see figure 2 for difference images of slice 6, phase 14).

From figure 2, it can be seen that the registration has successfully corrected the stacks of the misaligned image, producing an anatomically more correct image within the region of the heart. To confirm this, the sum of squared differences (SSD) of the motion-free SA image and the misaligned SA image before and after registration was calculated for each of the 25 phases within the area of the heart. The SSD of the motion-free SA image and the misaligned SA image after registration is, on average, 43.5% less than the SSD of the motion-free SA image and the misaligned SA image before registration. This implies that the misaligned SA image after registration is more similar than the misaligned SA image before registration when compared to the motion-free SA image.

5.3 Left ventricular volume analysis

We assessed the accuracy of ventricular function measures calculated from the misaligned slices, and from the same slices after slice-to-volume registration. In both cases, we compared their result to the ventricular function measurements from the motion-free acquisition. On average, the misaligned volume deviates from the motion-free volume by 3.2ml whereas the realigned volume only deviates by 1.5ml. This is further confirmation that the slice-to-volume registration has sufficiently recovered the misalignment due to differing breath hold positions.

6 Conclusions and discussion

We have shown that the accuracy of slice-to-volume registration is sufficient to use this method to realign misaligned multi-slice SA cardiac stacks for improved visualization, as seen in section 5.2, and more accurate LV volume analysis, as shown in section 5.3. Furthermore, we have shown, with the use of a simulation experiment that the registration algorithm is sufficiently robust to recover typical movements expected in misaligned slice stack positions. Such improvements will be beneficial for diagnosis and treatment planning for historic data as well as for data acquired when the patient is unable to hold their breath at a consistent position.

At present, the 3D SSFP volume is also acquired in multiple breath holds. It is therefore vital that the breath hold positions of this image are consistent. An expert cardiologist verified the 3D volume that we have used in these experiments to be well aligned by visual inspection. Inconsistencies in the alignment of the volume could have dramatic effects on the accuracy of the slice-to-volume registration algorithm. For this reason, future work will investigate whether similar registration accuracy and robustness can be obtained when the 3D SSFP volume is replaced by a 3D volume which can be acquired in a single breath hold, such as a 3D k-t BLAST volume [6].

Acknowledgements

A.G. Chandler is funded by Philips Research Laboratories, Hamburg, Germany. J.A. Schnabel has received funding from Philips Medical Systems, Best, The Netherlands. The authors would like to thank Dr. Maxime Sermesant for supplying software for the visualizations in figure 1, and Dr. Vivek Muthurangu for his expert opinion of the images.

References

1. Bellenger NG, Marcus NJ, Rajappen K, *et al.*, "Comparison of techniques for the measurement of left ventricular function following cardiac transplantation", *Journal of Cardiovascular Magnetic Resonance*, **4**, 255-263, 2002.
2. Ehman RL, McNamara MT, Pallack M, *et al.*, "Magnetic resonance imaging with respiratory gating: Techniques and advantages", *Amer. J. Roentgenol*,**143**, 1175-1182, 1984.
3. Sachs T, Meyer CH, Pauly JM, *et al.*, "The real-time interactive 3D-DVA for robust coronary MRA", *IEEE Transaction on Medical Imaging*, **19**, 73-79, 2000.
4. Nguyen TD, Wang Y, and Prince MR, "A novel navigator technique for fast and direct detection of 3D displacement of the coronary arteries", *In Proceedings of 9th Annual Meeting ISMRM*, 44-52, 2001.
5. McLeish K, Hill DLG, Atkinson D, *et al.*, "A study of the motion and deformation of the heart due to respiration", *IEEE Transaction on Medical Imaging*, **21**, 1142-1150, 2002.
6. Kozerke S, Tsao J, Razavi R, *et al.*, "Accelerated Cardiac Cine 3D Imaging using k-t BLAST", *Magnetic Resonance in Medicine*, 2004 (in press).
7. Studholme C, Hill DLG, and Hawkes DJ, "An overlap invariant entropy measure of 3D medical image alignment", *Pattern Recognition*, **32**, 71-86, 1999.

A Coupled Active Contour Model for Myocardial Tracking in Contrast Echocardiography

Norberto Malpica[a], María Ledesma-Carbayo[a] , Andrés Santos[a], Esther Pérez[b], M.Angel García-Fernandez[b] and Manuel Desco[b]

[a]Departamento de Ingeniería Electrónica, E.T.S.I. Telecomunicación, Ciudad Universitaria s/n 28040 Madrid Spain, [b] Laboratorio de Imagen, Medicina Experimental, Hospital G. Gregorio Marañón, Madrid

Abstract. Contrast echocardiography has been proposed as an indicator of myocardial perfusion in a non-invasive way. Reperfusion curves can be obtained by destroying all the microbubbles using an ultrasound pulse with high mechanical index and acquiring images during the reperfusion process. Quantitative parameters describing the process can be obtained from the curves. To analyze the complete myocardium, we propose a method for the simultaneous segmentation and tracking of endocardium and epicardium in myocardial contrast echocardiography sequences. The model consists of two active contours, guided by optical flow estimates. The evolutions of the two contours are coupled geometrically using a novel scheme that imposes stability in wall thickness, to deal with low contrast regions in the epicardial contour. Both a closed and an open model have been designed, to account for the different acquisition views used routinely. The model has been evaluated with experimental and clinical sequences, comparing the results with manual segmentations carried out by an expert.

1 Introduction

Myocardial contrast echocardiography (MCE), due to its ability to assess microvascular integrity, has been shown to provide markers of successful reperfusion of acute myocardial infarction [1]. Viability of the myocardium is estimated by the degree of myocardial opacification following contrast injection. A method for obtaining quantitative parameters of the reperfusion process consists in destroying the contrast microbubbles with an ultrasound pulse of high energy (high mechanical index) and acquiring images continuously during the reperfusion process. It is thus possible to obtain the wash-in curve showing the refilling of the region after the destruction of the microbubbles [2]. The segmentation and time tracking of the complete myocardium in sequences acquired with a real-time acquisition would allow to analyze simultaneously wall motion and perfusion of all myocardial segments. To achieve this task, Caiani et al. [3] have proposed a method for segmenting each frame separately. The endocardium is segmented interactively in every frame of the sequence and the epicardium is obtained by dilating the edge of the endocardium a fixed width. Garcia et al. [4] have evaluated a snake model combined with an active shape model to segment four-chamber views. A previous segmentation of a high number of images is required to train the model, with the intrinsic drawback that the model allows to segment only sequences acquired in a specific view. In this work we propose the use of two active contours related by a novel coupling scheme to simultaneously segment endocardium and epicardium. A motion estimation step is incorporated to track the contours between frames.

2 Active contour model

Each active contour is represented as a discrete set of points or snaxels $\{v_1,v_2,... v_n\}$ with $v_i = (x_i,y_i)$ [5], and the energy of the contour is defined as:

$$\varepsilon = \int_0^1 (\varepsilon_{\text{int}}(v(s)) + \varepsilon_{ima}(v(s)) + \varepsilon_{const}(v(s)))\, ds \quad (1)$$

where ε_{int} is the internal energy, dependent on the shape of the contour, ε_{ima} is the image dependent energy, and ε_{const} represents the external constraints imposed by the user.

Two coupled active contours are used in our model, to segment the endocardium and the epicardium, respectively. The evolutions of the two curves are coupled using geometric constraints. To segment a complete sequence, only an initialization of both curves in the first frame is required. The user marks several points, and the initial contours are interpolated from these points using a B-spline curve. An evolution of the curves is then carried out to adjust the curves to image gradients. In the remaining frames, segmentation is carried out in two

stages. To compensate inter-frame displacement a first evolution is performed, based on motion estimation. A second evolution is then performed on each snake using the gradient as image energy, with distance restrictions between snakes to correct the evolution on regions of the image with low gradient.

To be able to deal with sequences acquired either in short axis views or in apical two-chamber and four-chamber views, we have designed a closed model and an open one. In the closed model the first and last points are linked together. In the open model, the first and last points of the curve only have one neighbour each. Definition of curvature and distance energies in these cases is not straightforward [5]. We set the curvature energy, $\varepsilon_{curvature} = 0$ in both ends, so the end points tend to follow their neighbours.

2.1 Internal energy

In our model, the internal energy is the addition of two terms. The first one is a curvature energy, based on the second derivative. The second term is an energy aimed at distributing snaxels uniformly along the contour. The external energy is based on the gradient of a diffusion-filtered version of the image. We have used a simple gradient based on finite differences. The energy is computed as the inverse of the gradient value.

2.2 Distance constraints

Some parts of the epicardium show a low gradient or no gradient at all due to blooming or other artefacts in the acquisition process. On these regions, an evolution of the curve based only on the gradient of the image would lead to incorrect results. The endocardium, on the contrary, is usually well depicted due to the high intensity of contrast in the cavity. We assume that the segmentation of the endocardium based only on the gradient is correct, and introduce a distance constraint, modeled as two coupling energies between the curves.

To implement this coupling, we define two distance-based energies, denoted as hard distance constraint and soft distance constraint. The hard constraint, applied to both inner and outer curves, imposes a maximum and a minimum distance between both curves, similar to that proposed by Zeng [6] for segmenting the brain cortex. When the distance between the curves is outside the allowed distance range, the energy has a very high constant value. The distance was defined based on previous manual segmentations and similar constraints proposed in the literature.

The soft constraint, applied only to the evolution of the outer curve, is a novel coupling scheme. It aims at preserving the mean distance between curves in regions with low gradients. A weighted mean distance between the curves is computed, weighting the distance of each snaxel by the value of the image gradient at that point:

$$\bar{d}_i = \frac{1}{N} \sum_{k=1}^{N} g_k \; d(v_k, C_j) \tag{2}$$

The energy of every snaxel is then computed as follows:

$$\varepsilon_{dist} = d(v_i, C_j) - \bar{d}_i \tag{3}$$

This distance energy is again weighted by the value of the gradient at every snaxel, so that the contribution of this energy is low in points with a high gradient value.

Figure 1 shows the effect of the soft constraint based on mean distance. On the left, the result of the evolution without this constraint is shown. On the right, the result with the constraint is presented. As can be observed, sections of the curve with a high gradient are not affected, while the result is corrected in sections with gradient dropouts.

2.3 Motion energy

To take advantage of temporal information and make the segmentation more efficient, we define a motion energy based on inter-frame displacement estimation. As reported in [7] we tested several optical flow methods to track regions of interests in contrast echocardiography sequences. The best results were obtained with the method proposed by Lucas and Kanade [8].

A multiscale version of the algorithm was designed, using a four-level Gaussian pyramid to capture large displacements without a major increase in computational load. The displacement is first computed at the lowest resolution. The algorithm is then applied at higher levels, starting from the displacements obtained in the previous level.

Once we have obtained a displacement for every control point, the energy at each point of the neighbourhood is defined as the distance in pixels to the displaced point. Internal energies are also taken into account. The compete model is thus:

$$\varepsilon = \int_0^1 (\alpha \varepsilon_{mot}(v(s)) + \beta \varepsilon_{\text{int}}(v(s)) + \gamma \varepsilon_{const}(v(s)))\, ds \tag{4}$$

where ε_{mot} represents motion energy. In all experiments, the parameters used were $\alpha = 1.0$, $\beta = 0.6$, $\gamma = 0.8$.



Fig 1. Result of curve evolution without distance constraints (left) and including the constraints (right)

3 Experiments and Results

For evaluating the closed model, six short axis sequences acquired from pigs during experimental surgery were used, with a total of 221 images. The open model was tested on 181 images from 6 sequences, obtained in clinical routine from 6 different patients. Three of the sequences were two chamber views and three were four chamber views. Images were acquired with Contrast Pulse Sequencing (CPS), a real-time acquisition method with an Acuson Sequoia (Acuson-Siemens) scanner. All the images were manually segmented by an expert observer, and by the automatic model. To compare both segmentations, two parameters were obtained: mean distance between curves and Hausdorff distance (maximum distance from a point of a curve to the other curve). For the close model, the degree of overlap between segmented surfaces was also computed. The results are summarized in table 1 (closed model) and table 2 (open model). Examples of the segmentation of two frames are shown in figure 2.

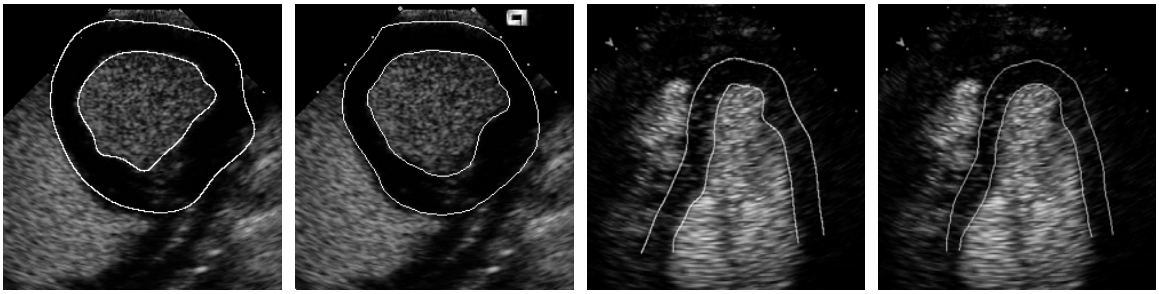
The proposed algorithm achieved good results both in endocardium and epicardium. Differences with manual segmentation results are larger in the epicardium, due to the lack of information in that part of the image. Results are similar to those presented by [4], but their model required a priori training by manual segmentation of a high number of images.

	Degree of overlap	Mean distance (mm)	Hausdorff distance (mm)
Endocardium	97.7 % ± 0.12 %	1.46 ± 0.70	2.34 ± 1.54
Epicardium	94.5 % ± 2.11 %	1.87 ± 0.63	3.21 ± 1.97

Table 1. Comparison between manual and automatic segmentation results for the closed model

	Mean distance (mm)	Hausdorff distance (mm)
Endocardium	1.33 ± 0.69	3.23 ± 2.01
Epicardium	1.75 ± 0.82	3.91 ± 2.52

Table 2. Comparison between manual and automatic segmentation results for the open model



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

Figure 2. Comparison between automatic and manual segmentations. Automatic (a) and manual segmentation (b) of a short axis frame. Automatic (c) and manual segmentation (d) of a four chamber view frame.,

4 Discussion and Conclusion

We have proposed a model consisting of two coupled active contours. Our model is automatic, except for the definition of the myocardium by the user in the first image. The advantage with respect to previous proposals is that it does not require a training of the model, avoiding the need of time-consuming expert interaction and making it more general, allowing to segment sequences acquired in different views. Another advantage is the inclusion of an optical flow stage to compensate for motion between frames, which allows for an efficient tracking even in sequences acquired with a low frame rate. When there is no other information (lack of image contrast), the method makes the assumption of uniform myocardial thickness in each frame. This may not be correct in some pathological cases, in which some segments may have a reduced thickness. No other a priori information is needed.

References

1. M. A. Vannan and B. Kuersten, "Imaging techniques for myocardial contrast echocardiography," *European Journal of Echocardiography*, vol. 1, pp. 224-226, 2000.
2. K. Wei, A. R. Jayaweera, S. Firoozan, A. Linka, D. M. Skyba, and S. Kaul, "Quantification of myocardial blood flow with ultrasound-induced destruction of microbubbles administered as a constant venous infusion," *Circulation*, vol. 97, pp. 473-83, 1998.
3. E. G. Caiani, R. M. Lang, S. Caslini, K. A. Collins, C. E. Korcarz, and V. Mor-Avi, "Quantification of regional myocardial perfusion using semiautomated translation-free analysis of contrast-enhanced power modulation images," *Journal of the American Society of Echocardiography*, vol. 16, pp. 116-23, 2003.
4. J. Garcia, D. Rotger, P. Radeva, F. Carreras, and R. Leta, "Contrast Echography Segmentation and Tracking by a Trained Deformable Model," *Proceedings of Computers in Cardiology*, Thessaloniki, 30:173-177 Thessaloniki Greece, 2003.
5. S. Lobregt and M. A. Viergever, "A discrete dynamic contour model," *IEEE Transactions on Medical Imaging*, vol. 14, pp. 12-24, 1995.
6. X. Zeng, L. H. Staib, R. T. Schultz, and J. S. Duncan, "Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 927-37, 1999.
7. N. Malpica, A. Santos, M. A. Zuluaga, M. J. Ledesma, E. Pérez, M. A. García-Fernández, and M. Desco, "Tracking of Regions-of-Interest in Myocardial Contrast Echocardiography," *Ultrasound in Medicine and Biology*, vol. 30, pp. 303-309, 2004.
8. B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proceedings DARPA IU Workshop*, 1981.

Detection of Perfusion Defects using a Spatio-Temporal Analysis of Contrast Ultrasound Image Sequences

Quentin R Williams[a] and J Alison Noble[a]

[a]Medical Vision Laboratory, University of Oxford, {quentin, noble}@robots.ox.ac.uk

Abstract. Tissue perfusion imaging is becoming an increasingly employed method to assess internal organ blood supply and flow in clinical applications. Reports on the clinical implementation of contrast echocardiography for perfusion imaging and quantitative measurement have increased dramatically of late, due to the accessibility and non-invasiveness of ultrasound. This paper describes a method that identifies different areas of perfusion in a contrast ultrasound perfusion study, providing regionalised maps of tissue perfusion. Pixels in an image sequence are automatically classified into different classes, using a Bayesian Factor Analysis Model set in a Markov Random Field framework; utilising both the temporal and spatial characteristics of the pixels for classification. The purpose of this study was to implement and evaluate this method in a myocardial perfusion clinical setup. Results on multiple in-vivo datasets are presented indicating good agreement with expert analysis and angiogram tests, showing how this method can be used to identify normal and abnormal perfused tissue.

1 Introduction

The evaluation of tissue perfusion in various parenchymatous organs is important in the diagnosis and determination of the severity of ischemic disease. The ability to measure internal organ blood supply and flow in-vivo allows the assessment of tissue viability and the analysis of disease state in various organs like the heart, liver and kidneys. Tissue perfusion has been traditionally analysed by nuclear medicine imaging procedures, like T1-SPECT, which measure cell membrane integrity, or more recently by PET-FDG, which shows metabolism and blood flow rates, but limitations of these techniques include low spatial resolution and the use of ionising radiation.

The availability, low-cost and non-invasiveness of contrast-echocardiography have fostered an increasing interest in the use of this modality to provide an accurate diagnosis of tissue perfusion. Both quantitative and qualitative measurements of tissue perfusion can be made by injecting a contrast agent (microbubbles) intravenously and then imaging the changes in signal intensity as the contrast agent makes its pass through the organ. The contrast agents submerged in blood, increase the echo-backscatter of perfused tissues and blood pool in cavities, and are ideal as they remain intravascular and have a particle size similar to red blood cells [1]. This allows for better visualisation of blood flow information and regional perfusion. While doctors have so far relied on subjective visual inspection, techniques have been developed to extract information from the contrast-enhanced studies. These are mostly based on the 'negative bolus indicator dilution technique' introduced by Wei et al. [2]. After a high power ultrasound induced destruction of microbubbles; their replenishment is assessed during a constant venous infusion of a contrast agent. The replenishment (wash-in) curves, showing the refilling of microbubbles, are then fitted to an exponential function: $y = A(1-e^{-\beta t})$, to extract parameters involving the mean microbubble velocity (β) and tissue blood volume (A). The quantification is then used to assist the physician in his/her interpretation of the study. Other semi-quantitative ROI approaches include depicting the parameters on a pixel-by-pixel basis [5]; using frequency domain analysis to quantify blood flow and transit time [3]; or the use of a calibration method with mean pixel intensities to differentiate tissue segments into akinetic and kinetic regions [4]. These pixel-based techniques all suffer from ad hoc smoothing in space and time, and the loss of temporal information that is available by analysing the correlation between pixels.

The purpose of this paper is to show how a statistical method can be used to identify different areas of perfusion in a contrast ultrasound perfusion study. Pixels in an image sequence are automatically classified into different classes, by analysing their distinct temporal relationships. Using a Bayesian Factor Analysis (BFA) model and incorporating spatial information through a Markov Random Field, different regions of tissue perfusion are identified (whether normal or abnormal) and the nature and structure of the tissue perfusion can be examined. This can aid in the recognition of tissue regions with abnormal blood flow as well as its spatial extent, as is shown with a few clinical examples. This BFA-MRF algorithm will be presented at the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) 2004. In this paper however the method have been extended to allow for better initialisation and tested on clinical datasets. The BFA-MRF algorithm is summarised in section 2, with the results presented and discussed in sections 3 and 4.

2 Methods

2.1 Bayesian Factor Analysis using a Markov Random Field prior model

A novel spatio-temporal technique is presented to assess tissue perfusion by automatically classifying the ultrasound images into different regions of perfusion. This is done in a global manner by analysing the temporal pattern of relationships between pixels, using a Bayesian Factor Analysis model, and incorporating spatial information through a Markov Random Field. Briefly, this method treats the classification as a statistical problem, which involves assigning to each pixel a class label taking a value from the set $L = \{1,2,\ldots,l\}$, where each pixel is indexed by a two-dimensional rectangular lattice $S = \{1,2,\ldots,n\}$ and characterised by a p-variate vector of intensity values $\mathbf{y}_i = (y_{i1},\ldots, y_{ip})$, $i \in S$. In this case each observation vector $\mathbf{y}_i$ represents an intensity-time curve for a single pixel location. The problem of classification is then to estimate the true but unknown labeling configuration, $\mathbf{x}^*$, given the observed intensity time vectors, $\mathbf{Y}' = (\mathbf{y}_1,\ldots, \mathbf{y}_n)$. In particular, the maximum a posteriori (MAP) estimate of $\mathbf{x}$ is used:

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in X} \{P(\mathbf{Y} \mid \mathbf{x})P(\mathbf{x})\}. \tag{1}$$

The right-hand side of the above equation contains two parts: $P(\mathbf{Y} \mid \mathbf{x})$ and $P(\mathbf{x})$, which are defined as a Bayesian Factor Analysis likelihood distribution and a Markov Random Field prior distribution, respectively. What remains is the estimation of the parameters of these two distribution functions, where the BFA model is constructed as a generative latent variable model,

$$(\mathbf{y}_i \mid \mu, \Lambda, \mathbf{f}_i) = \mu + \Lambda\mathbf{f}_i + \varepsilon_i , \tag{2}$$

for each observation vector $\mathbf{y}_i$ (i=1,…,n), where μ is the overall population mean, Λ is a matrix of constants called the factor loading matrix; $\mathbf{f}_i = (f_{i1},\ldots,f_{il})$, $l \in L$, is the factor score vector for pixel i; and the ϵ_i's are assumed to be mutually uncorrelated and Normally distributed $N(0,\Psi)$ variables. Since the parameters μ, Λ, the $\mathbf{f}_i$'s, and Ψ are all unobservable, a Normal likelihood distribution for each $\mathbf{y}_i$ is assumed, and written as:

$$p(\mathbf{y}_i \mid \mu, \Lambda, \mathbf{f}_i, \Psi) = (2\pi)^{\frac{-p}{2}} \left|\Psi\right|^{\frac{-1}{2}} e^{-\frac{1}{2}(\mathbf{y}_i-\mu-\Lambda\mathbf{f}_i)'\Psi^{-1}(\mathbf{y}_i-\mu-\Lambda\mathbf{f}_i)} . \tag{3}$$

The probability of an MRF realisation, $\mathbf{x}$, is given by the Gibbs distribution:

$$P(\mathbf{x}) = Z^{-1} e^{(-\omega U(\mathbf{x}))} , \tag{4}$$

where

$$U(\mathbf{x}) = \sum_{c\in C} V_c(\mathbf{x}) , \tag{5}$$

is the energy function which is a sum of *clique potentials* $V_c(\mathbf{x})$ over all possible cliques $\boldsymbol{C}$. Z is a normalisation term and ω is a positive constant which controls the size of clustering

The novelty of this particular algorithm stems from the way it interlinks the factor scores in the BFA model to the prior probability of the MRF model. The factor score vector indicates how much an observation belongs to a particular class. It is therefore assumed that the prior probability of the factor scores matrix, $\mathbf{F}' = (\mathbf{f}_1,\ldots, \mathbf{f}_n)$, follows the same prior probability of the classification configuration, $\mathbf{x}$, and in fact that the factor score for each hidden factor (or class) is equivalent to the posterior probability of the class label. For every $l \in L$ and $i \in S$

$$f_{il} = P(\mathbf{y}_i \mid l)P(x_i = l) . \tag{6}$$

Using the prior probability and the likelihood function with respect to x_i and f_{il} gives

$$f_{il} = Z^{-1} e^{(-\omega U(x_i))} \times (2\pi)^{\frac{-p}{2}} \left|\Psi\right|^{\frac{-1}{2}} e^{-\frac{1}{2}(\mathbf{y}_i-\mu-\Lambda_l f_{il})'\Psi^{-1}(\mathbf{y}_i-\mu-\Lambda_l f_{il})} . \tag{7}$$

Therefore the posterior probability values obtained through the MRF-MAP classification can directly be used as the factor scores. Thus, the strategy underlying this algorithm can be summarised as follows: (1) Estimate the labelling configuration, $\hat{\mathbf{x}}$, using the current estimate of the parameters; (2) use it to specify the factor scores

matrix, **F**; (3) and then estimate the new values for the parameters, μ, Λ, and Ψ. These steps are iteratively repeated until suitable convergence is reached.

2.2 Experimental Methodology

The BFA-MRF algorithm was initiated with a tree-structure K-means (TSKM) initial segmentation and implemented in a multi-scale framework to improve convergence. The method was tested on clinical datasets all obtained during a constant venous infusion of the SonoVue® (Bracco International B-V) contrast agent. The ultrasound was imaged with the Power Pulse Inversion technique described in [1], and image sequences were ECG-triggered to keep only the end-systolic (end of T-wave) frames. Three datasets were used in the study: (1) A 2-chamber apical view of a healthy patient; (2) the rest and stress images in a 4-chamber apical view of a patient with a reversible apical defect as well as (3) the 3-chamber apical view of the same patient. The datasets were all graded as "good" quality by a cardiologist, and each image sequence consisted of 13 gated frames. A colour-coded classification image, (see Figure 1) along with the mean intensity-time profiles of each class was produced and compared with visual analysis from an expert. In the case of datasets (2) and (3) above, the results were also compared with analysis from an angiogram test. The parametric method described in [2] was applied to the 3-chamber view and compared with results from the BFA-MRF algorithm.

3 Results

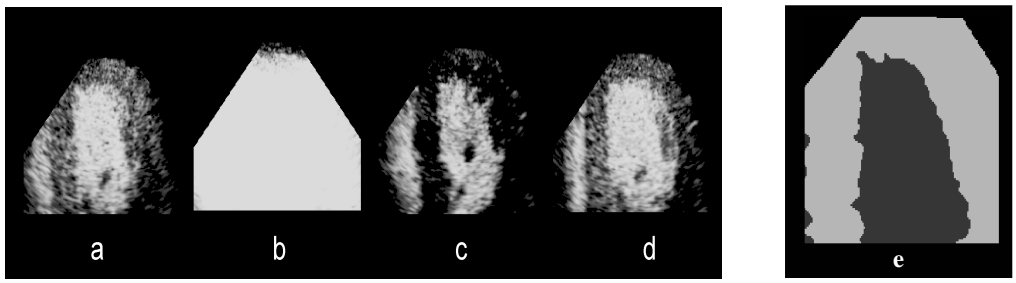


Figure 1. Dataset of healthy volunteer: a) Constant venous infusion of contrast agent. b) high power frame to destroy the microbubbles. c) and d) 1st and 5th frame after destruction (low power). e) Classification obtained showing 2 different types of perfusion present.

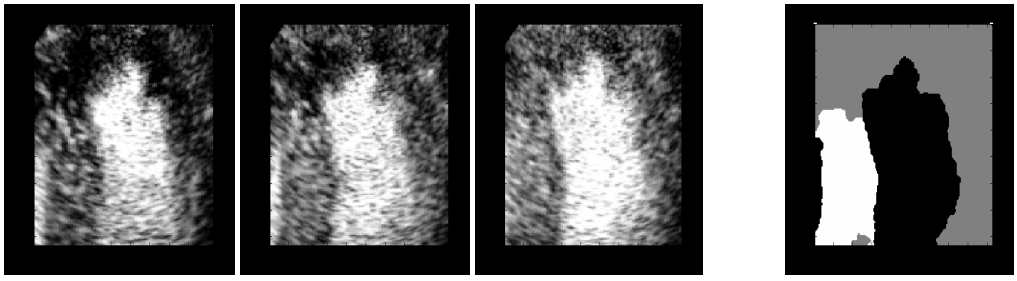
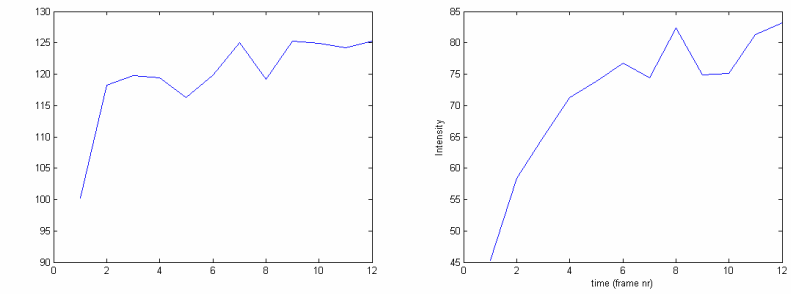


Figure 2. Frames 1,3 and 12 of the 4-chamber stress image sequence and the classification result

Figure 1 shows four frames taken from the dataset of the healthy volunteer depicting the replenishment of the contrast agents in subsequent frames. The classification map indicates that there are two perfusion classes present; (1) the whole myocardium operating with normal perfusion (grey) and (2) The cavity (black) which stays at a constant intensity level throughout the sequence. Figure 2 shows the 1st, 3rd and 12th frames of the 4 chamber stress image sequence, as well as the classification result. Three distinct classes were found: (1) The cavity (black), (2) a region of normal perfusion (white) and (3) abnormal perfusion (grey). Normal and abnormal perfusion are distinguished using the mean intensity-time profiles, as can be seen in Figure 3. In the abnormal case it takes longer to reach its final intensity value, indicating a lower blood velocity, while the final intensity is also lower than in the normal case, resulting from an inferior blood volume in the abnormal myocardium segment. These results agreed well with the angiogram of the patient which showed a subtotal occluded branch of the Left anterior descending coronary artery (LAD) as seen in the apex of the classification image, and a branch of the Left Circumflex Coronary Artery (LCX) which had a 95% stenosis, as seen in the lateral wall of the classification image. In all of the cases the classification obtained from the BFA-MRF method agreed with visual inspection from an expert; as well as analysis using the parametric method and the angiogram tests.



(a) (b)

Figure 3. Mean intensity-time profiles of (a) Normal segment and (b) Abnormal segment for the 4-chamber view (note the different scales on the y-axis)

4 Discussion

This paper has shown how both the spatial and temporal characteristics of the pixels can be incorporated into a single statistical model for classification, which automatically identifies different types of perfusion. By plotting the mean intensity-time profiles of each class, showing the flow of the contrast agent through the section; alongside the colour-coded classification map, it was shown that the method can clearly distinguish between normal and abnormal perfused regions in the myocardium. In this manner regionalised maps of tissue perfusion are provided showing the spatial extent of each region. Other dynamic parameters, like the blood velocity and blood volume can also be obtained from the intensity-time profiles if necessary. The ultimate goal for contrast perfusion imaging is to define the area of tissue infarction and perfusion deficit, and to identify any ischemic tissue that can be salvaged by medical or surgical therapy. Although the BFA-MRF method has only been tested on a small number of cases, the preliminary results on in-vivo data are encouraging, showing agreement with expert analysis as well as other current methods in use. It is important to note that all of the data used in this paper were of "good" quality. A validation study testing the algorithm over a range of clinical datasets is currently underway which should provide more rigorous analysis of the clinical practicability of this method. Other issues to consider are alignment of the images within a sequence which should improve classification results. In the examples in this paper the images were ECG-triggered, but this alone does not completely remove misalignment errors and an extra image-registration step might be considered.

Easy and accessible, contrast echocardiography perfusion imaging is more appealing for routine clinical examinations than nuclear imaging techniques. Research regarding its accuracy and clinical viability is still needed, but early work has shown that this technique can aid the clinician in the diagnosis and assessment of diseased tissues.

Acknowledgements

The authors would like to thank Drs. Alex Ehlgen and Harold Becher from the John Radcliffe Hospital, Oxford, for their invaluable help in providing the datasets and discussions regarding the results.

References

1. H. Becher & P. Burns. *Handbook of Contrast Echocardiography.* Springer Verlag, 2000.
2. K. Wei, et al. "Quantification of myocardial blood flow with US induced destruction of microbubbles administered as a constant venous infusion". *Circulation*, vol. 97:473-83, 1998.
3. V. Mor-Avi, S. Akselrod, D. David, L. Keselbrener, and Y. Bitton. "Myocardial transit time of the echocardiographic contrast media", *Ultrasound Med Biol*, vol. 19:635-48, 1993.
4. A. Yano, et al. "Myocardial Contrast Echocardiography with a new Calibration Method Can Estimate Myocardial Viability in Patients With Myocardial Infarction.", *J. of the American College of Cardiology*, vol. 43:1799-86, 2004.
5. A.Z. Linka, J. Sklenar, K. Wei, et al. "Assessment of transmural distribution of myocardial perfusion with contrast echo". *Circulation*, vol. 18:1912-20, 1998

Unsupervised Classification of Late Gadolinium Enhancement Cardiovascular MR

Marcus Ellington , Xiao-Peng Hu, Anna John, Dudley Pennell, Guang-Zhong Yang

Royal Society/Wolfson Foundation MIC Laboratory
Imperial College of Science, Technology and Medicine, London SW7 2BZ, U.K.

Abstract. Patients suffering from ischemic left ventricular dysfunction but with viable myocardium may recover contractile function after revascularisation. The accurate identification of viable myocardium is important for guiding the subsequent treatment and establishing its prognosis. Late enhancement CMR with Gd-DTPA can depict necrotic or fibrosed areas after myocardial infarction, permitting discrimination between subendocardial and transmural defects [1]. This paper presents an automatic method for segmenting infarcted and non-infarcted regions with 3D multi-slice late gadolinium enhancement CMR. Validation of the method was performed on data acquired from 10 patients, showing similar delineation accuracy to that of an experienced observer, but having the advantage of no inter- and intra-observer variability.

1 Introduction

Restricted coronary blood supply can lead to left ventricular dysfunction in patients with ischemic heart disease. Cell death occurring within the affected myocardium can further cause irreversible injury, thus resulting in myocardial infarction. Otherwise, the myocardium may remain in a viable state, from which it can recover its contractile function after revascularisation. For the management of ischaemic heart disease, it is important to accurately distinguish between viable and infarcted regions of myocardium. There is evidence that nonviable myocardium should not be revascularised, as this will not improve systolic function, and may pose a greater mortality risk than alternative therapy [2,3].

Late enhancement MR is a recent technique that has the potential of accurately identifying nonviable regions of the myocardium. Supporting evidence from animal studies shows that the size and shape of the regions exhibiting hyper-enhancement correspond well to that of necrotic tissue identified in histological images [1,4]. The ability to assess the extent of infarction in late enhancement images has been shown to have promising clinical applications, particularly in differentiating the underlying causes of heart failure [5] and performing clinical risk assessment [6]. It may also be possible to combine the assessment of late enhancement with cine MRI and first-pass myocardial perfusion imaging to accurately differentiate between regions of healthy myocardium, injured but viable myocardium and infarcted tissue, making way for further clinical applications.

Thus far, existing studies have mainly relied on the quantification of the degree of hyper-enhancement by performing manual delineation. These methods can be time consuming and usually result in significant inter-observer variability. An automatic method is therefore of great importance to obtain accurate and consistent quantification of late enhancement. Current research based on supervised classification such as support vector machines [7] has given encouraging classification results, the purpose of this paper is to provide an unsupervised alternative that is simple and efficient to use. Validation of the method is performed on data acquired from 10 patients, demonstrating similar delineation accuracy to that of an experienced observer.

2 Method

2.1 Image Acquisition

For this study, 10 patients suffering from ischemic heart disease with previous myocardial infarction were recruited with informed consent. All subjects were imaged in supine position on a Siemens Sonata 1.5T scanner (Siemens, Erlangen, Germany). A bolus of gadolinium-DTPA 0.1 mmol/kg (Magnevist, Schering, Berlin, Germany) was given intravenously. Fifteen minutes later, late enhancement images were obtained as follows. A standard ECG-gated FLASH-IR sequence was used with a slice thickness of 8 mm, 2mm gap. The inversion time between 400 and 480 ms was chosen to null non-enhancing myocardium. Imaging planes were chosen along the

two long axes of the heart (fourchamber and twochamber). A stack of short axis images, typically with 7-9 slices, was acquired spanning the ventricles from atrio-ventricular grove to the apex.

2.2 Region Classification

With late enhancement MR, due to the time delay after the intravenous administration of Gd-DTPA, the agent is washed out of the healthy myocardium and partially filtered out of the bloodstream passing through the liver. From the acquired images the infarcted regions may be distinguished by their delayed but sustained enhancement relative to the healthy myocardium. This continued enhancement is due to the reduced rate at which blood diffuses to and from the necrotic or fibrosed myocardial tissue. In order to segment the hyper-enhanced myocardial tissue, we make use of the intensity histogram to automatically determine an optimal threshold. The intensity distribution within the myocardium will exhibit two distinct clusters, one at a lower level for the viable myocardium and another at high intensity for the areas of infarct. To model this, the Finite Gaussian Mixture Model [8] is used.

Let us denote each component of a Gaussian mixture model as $g_i(x/\theta_i)$ where θ_i represents the parameters of the component g_i and x is the intensity value. Let there be a total of G components in the model. The Gaussian mixture density can be defined as:

$$pg(x/c,\theta) = \sum_{i=1}^{G} c_i g_i(x/\theta_i) \tag{1}$$

where $c = (c_1, c_2, ...c_G)$ is a vector of mixing weights, such that:

$$\sum_{i=1}^{G} c_i = 1 \tag{2}$$

The Gaussian component is defined as:

$$g_i(x/\theta_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2} \right) \tag{3}$$

where $\theta_i = (\mu_i, \sigma_i)$, μ_i is the mean and σ_i is the standard deviation. In our study, the number of Gaussian components is 2, *i.e.*, $G = 2$.

Using the intensity histogram $h(x)$ calculated from an image, the parameters c and θ of the Gaussian mixture can be determined by the following recurrent equations, which are derived from the EM algorithm [9].

$$p_i(x) = \frac{c_i^{old} g_i\left(x/\mu_i^{old}, \sigma_i^{old}\right)}{\sum_{k=1}^{G} c_k^{old} g_k\left(x/\mu_k^{old}, \sigma_k^{old}\right)} \tag{4}$$

$$c_i^{new} = \frac{\sum_x h(x) p_i(x)}{\sum_x h(x)} \tag{5}$$

$$\mu_i^{new} = \frac{\sum_x h(x) p_i(x) x}{\sum_x h(x) p_i(x)} \tag{6}$$

$$\sigma_i^{new} = \sqrt{\frac{\sum_x h(x) p_i(x) \left(x - \mu_i^{old}\right)^2}{\sum_x h(x) p_i(x)}} \tag{7}$$

With two Gaussian distributions for the representation of intensty values in normal and infarcted regions, we can determine the optimal threshold level for segmentation. In our study, the optimal threshold value is the intersection point of the two Gaussians, which is calculated analyticaly. In a final post-processing step, morphological operators [10] were introduced to eliminate isolated enhanced pixels.

3 Results

Validation of the proposed technique was performed on 79 short axis slices, obtained from the 10 subjects involved in this study. An experienced observer provided manual classifications, with a delineation of the myocardial borders and two separate delineations of the enhanced regions applied to each image. Figure 1 shows late enhancement images acquired from one of the patients studied, illustrating the regions of enhancement delineated manually by an expert and the corresponding automatically detected areas.

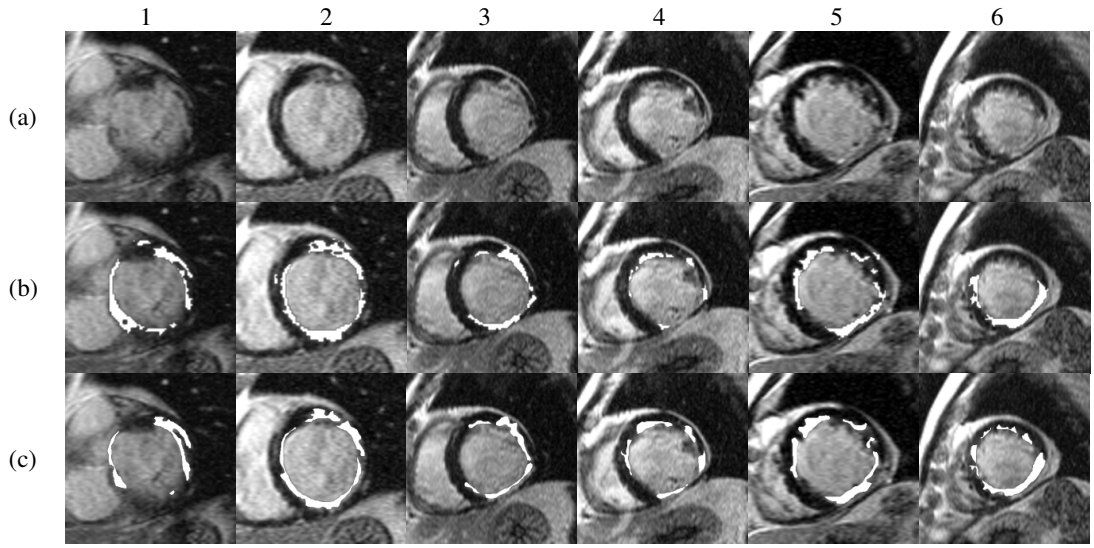
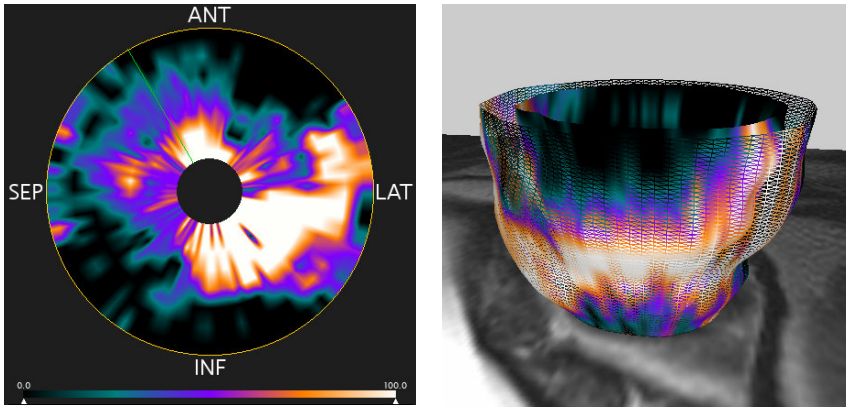


Figure 1. (a) The original acquired late enhancement images, (b) regions of automatically detected enhancement with the proposed technique, and (c) regions of enhancement delineated by an expert.

To assess overall correspondence of the manual and automatic delineation methods, the area of enhancement was calculated for each image, both as an absolute quantity and as a percentage of the total myocardial volume. The resulting measurements for the automatic method were compared to both of the manual classifications. Similar comparisons were also made between the two manual classifications, allowing the degree of intra-observer variability to be evaluated. These results are summarised in Table 1. To allow quantitative assessment, the percentage transmural hyper-enhancement is also represented in a bull's-eye plot, showing the 3D distribution of the affected myocardium from base to the apex, as shown in Figure 2. In this figure, the corresponding 3D visualisation is also provided as a reference.

	Manual 1 – Manual 2	Automatic – Manual 1	Automatic – Manual 2
Area of enhancement	0.944477	0.886718	0.911682
Percentage enhancement	0.938046	0.858825	0.904832

Table 1. Correlation ratio between manual and automatic delineation methods



(a) (b)

Figure 2. (a) Bull's-eye plot of the percentage transmural hyper-enhancement for the subject shown in Figure 1, and (b) the corresponding 3D visualisation of the affected regions superimposed on the epi- and endocardial surfaces.

4 Discussion and Conclusion

We have presented an automatic technique for the classification of late-enhancement gadolinium MR images. The proposed method has the advantage over other techniques in that the classification is fully automatic, and does not involve training data sets. Being an unsupervised classification method, the technique can be reliably generalised to data from different scanners with varying sequence parameters. The results obtained in this study indicate that the accuracy of the proposed automatic method is comparable to that of manual classification. It is worth noting that for this study, no explicit contextual information is incorporated into the classification process. It has been observed that in general the transmural progression of the infracted regions starts from endocardial segments, which can be incorporated into the current segmentation framework to further enhance the classification accuracy. Furthermore, for apical slices, use of the transverse imaging plane involves a large partial volume effect, and for reliable segmentation it is best to combine both short and long axis images. Nevertheless, the method presented here demonstrates a simple and practical way of obtaining reliable quantification of late-enhancement gadolinium MR.

References

1. Kim R.J., Fieno D.S., Parrish T.B., et al. *Relationship of MRI delayed contrast enhancement to irreversible injury, infarct age, and contractile function.* Circulation. 1999 Nov 9;100(19):1992-2002.
2. Allman K.C., Shaw L.J., Hachamovitch R., et al. *Myocardial viability testing and impact of revascularization on prognosis in patients with coronary artery disease and left ventricular dysfunction: a meta-analysis.* J Am Coll Cardiol 2002;39:1151–1158.
3. Hachamovitch R., Hayes S.W., Friedman J.D., Cohen I., Berman D.S. *Comparison of the short-term survival benefit associated with revascularization compared with medical therapy in patients with no prior coronary artery disease undergoing stress myocardial perfusion single photon emission computed tomography.* Circulation 2003;107:2900–2907.
4. Rehwald W. G., Fieno D.S., Chen E.-L., et al. *Myocardial Magnetic Resonance Imaging Contrast Agent Concentrations After Reversible and Irreversible Ischemic Injury.* Circulation, January 15, 2002; 105(2): 224 - 229.
5. McCrohon J.A., Moon J.C., Prasad S.K., et al. *Differentiation of heart failure related to dilated cardiomyopathy and coronary artery disease using gadolinium-enhanced cardiovascular magnetic resonance.* Circulation. 2003 Jul 8;108(1):54-9.
6. Moon J.C., McKenna W.J., McCrohon J.A., et al. *Toward Clinical Risk Assessment in Hypertrophic Cardiomyopathy With Gadolinium Cardiovascular Magnetic Resonance.* J Am Coll Cardiol. 2003 May 7;41(9):1561-7.
7. O'Donnell T.P., Xu N., Setser R.M., et al. *Semi-automatic segmentation of nonviable cardiac tissue using cine and delayed enhancement magnetic resonance images.* Proc. SPIE Int. Soc. Opt. Eng. 2003; 5031, 242
8. Redner, R.A. and Walker, H.F. *Mixture densities, maximum likelihood and the EM algorithm.* Society for Industrial and Applied Mathematics Review, 26, 2, (1984) 195-239.
9. Dempster, A.P., Laird, N.M. and Rubin, D.B. *Maximum likelihood from incomplete data via the EM algorithm.* Journal of the Royal Statistical Society, Series B, 39, (1977) 1-38.
10. Sonka M., Hlavac V. and Boyle R. *Image Processing, Analysis, and Machine Vision, 2nd Edition*, PWS Publishing, 1999, pp. 563-567.
11. Milligan G.W., Cooper M.C. *An examination of procedures for determining the number of clusters in a data set."* Psychometrika, vol. 50, no.2, pp. 159-179, 1985

A Partial Volume 3-D Gradient Magnitude Model

John Chiverton[1*], Kevin Wells[1] and Declan Johnson[2]

[1] School of Electronics and Physical Sciences, University of Surrey, Guildford, GU2 7XH,

[2] St. George's Hospital Medical School, Cranmer Terrace, London, SW17 0RE

Abstract. A feature space based on gradient magnitude and intensity information has previously been used to identify regions of pure material, Partial Volume (PV) voxels or boundary regions in 2-D images. However, biomedical images are often acquired with a 3-D acquisition process, for example MRI, PET and CT. This work therefore illustrates that a feature space based on 3-D gradient magnitude and intensity information improves the classification performance for 3-D imaging data over gradient magnitude methods that only use 2-D information. Equations that model the Probability Distribution Function (PDF) of 3-D gradient magnitude for volumetric 3-D imaging data are also presented.

1. Introduction

The Partial Volume (PV) effect is a well-known digital imaging artefact, typically associated with Magnetic Resonance Imaging (MRI) data [1,2]. In particular, neurological MRI data has been shown to contain a significant population of PV voxels [3]. A PV voxel can be described as a sample representing more than one tissue component due to the finite sampling resolution in the image acquisition process. Therefore image analysis methods that attempt to quantify a particular tissue component or material in an MRI data space should consider the PV effect. The performance of probabilistic classification algorithms can be improved with a feature space designed to separate the tissue component classes. It is often found that MRI tissue classes possess significant overlap in one-dimensional (1-D) intensity space. MRI data can also be acquired with multiple imaging sequences resulting in a higher dimensional feature space. Such multiple imaging sequences can produce better separation of the tissue component classes. However, these may not always be available and therefore alternative feature axes are necessary to provide better inter-class separation. Consequently, an easily obtained second feature is the gradient magnitude, which can be used to identify pixels in 2-D data with a high likelihood of belonging to an edge pixel (edgels) [4]. Concomitantly, non-edgels are usually modelled for the 2-D case with a Rayleigh PDF [5-9]:

$$p_{Rayleigh}(z) = \frac{z}{\sigma_{Ray}^2} \exp\left(-\frac{z^2}{2\sigma_{Ray}^2} \right). \tag{1}$$

Note, the Rayleigh distribution is equivalent to a chi-density for n=2 [10]. A second astute observation in [6,8], is that regions of higher gradient magnitude, such as regions of edgels in 2-D data (which will include PV voxels) can be modelled as a Rician distribution:

$$p_{Rician}(z) = \frac{z}{\sigma_{Ric}^2} \exp\left(-\frac{z^2 + \mu_{Ric}^2}{2\sigma_{Ric}^2} \right) I_0\left(\frac{z.\mu_{Ric}}{\sigma_{Ric}} \right). \tag{2}$$

PDFs (1) and (2) were then combined in [8], with Gaussian distributions (to model pure tissue intensity classes) and Gaussian distributions convolved with triangles (to model PV voxels) within a Bayesian classifier framework. The work presented here, extends the approach taken in [8] from modelling the 2-D gradient magnitude PDFs to modelling 3-D gradient magnitude PDFs in order to improve the classification performance for data acquired in 3-Ds.

2. Methodology

In common with the intensity model used in [8,11][1], pure tissue regions[2] for tissue class, j, are modelled as a Gaussian distribution, $p_G(g|i_j)$, with a mean, μ_j and a standard deviation σ_j. The intensity model for PV voxels is

*Electronic Mail: *j.chiverton@surrey.ac.uk*

[1] A similar intensity PV model was used in [2], but the intensity model in [8,11] provides a simpler method for separating individual classes.

[2] For neurological MRI data, pure tissue compartments usually include Grey Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF).

then modelled as a pure tissue Gaussian convolved with a triangular distribution, whose limits are defined by the means of the two tissues, *j* and *k*, under consideration. The triangular distribution, with slope, $M_{j,k}$, intercept, $C_{j,k}$, with normalisation factor, Q, is given by:

$$p_T\left(g \mid i_j, i_k\right) = \begin{cases} \frac{1}{Q}\left(M_{j,k} g + C_{j,k}\right) & \mu_j \le g \le \mu_k; \\ 0 & \text{elsewhere.} \end{cases} \tag{3}$$

The PDF arising from the Gaussian triangle convolution, $p_{GT}(g|i_j,i_k)$, is given by:

$$p_{GT}\left(g \mid i_j, i_k\right) = p_G\left(g \mid i_j\right) * p_T\left(g \mid i_j, i_k\right), \tag{4}$$

$$= -\frac{M_{j,k} g + C_{i,k}}{2.Q}\left\{\operatorname{erf}\left(\frac{g-\mu_k}{\sqrt{2\sigma_j^2}}\right) - \operatorname{erf}\left(\frac{g-\mu_j}{\sqrt{2\sigma_j^2}}\right)\right\} - \frac{M_{j,k}\sigma_j}{\sqrt{2\pi}.Q}\left\{\exp\left(\frac{g-\mu_k}{2\sigma_j^2}\right) - \exp\left(\frac{g-\mu_j}{2\sigma_j^2}\right)\right\}. \tag{5}$$

If the gradient magnitude is calculated using a simple 3-D gradient differential kernel, regions of pure tissue can be modelled by a Maxwell distribution [12], corresponding to a chi-density for n=3 [10]:

$$p_{Max}\left(z \mid i_j\right) = \sqrt{\left(2/\pi\right)}\frac{z^2}{\sigma_{Max,j}^3}\exp\left(-\frac{z^2}{2\sigma_{Max,j}^2}\right), \tag{7}$$

where $\sigma_{Max,j}$ corresponds to the standard deviation of the auto-correlation function of $p_G(g|i_j)$. For regions that are composed of voxels that belong to two single tissue classes, it can be shown that the gradient magnitude can be modelled with the generalised Rician PDF, for 3-D [12,13,14]:

$$p_{3DGM}\left(z \mid i_j, i_k\right) = \frac{z^{\frac{3}{2}}}{\sigma_{jk}^2\sqrt{\mu_{jk}}}\exp\left(-\frac{z^2+\mu_{jk}^2}{2\sigma_{jk}^2}\right).I_{\frac{1}{2}}\left(\frac{z.\mu_{jk}}{\sigma_{jk}^2}\right), \tag{8}$$

where σ_{jk} and μ_{jk} correspond to the standard deviation and means of the cross correlation function of the PDFs, $p_G(g|i_j)$ and $p_G(g|i_k)$. The gradient magnitude and intensity models can be combined into a feature vector, $\mathbf{x} = (g\ z)^T$, within a Bayesian framework [8], with a PV prior, $\varepsilon_{j,k}$, in order for an expression for the conditional PDF to be created:

$$p\left(\mathbf{x} \mid i_j\right) = \frac{1}{1+\sum_{\forall j\neq k}\varepsilon_{j,k}}\left\{p_G\left(g \mid i_j\right)p_{Max}\left(z \mid i_j\right) + \sum_{\forall k\neq j}\varepsilon_{j,k}.p_{GT}\left(g \mid i_j, i_k\right)p_{3DGM}\left(z \mid i_j, i_k\right)\right\}. \tag{9}$$

The performance of the new 3-D gradient magnitude PV model was compared with a 2-D gradient magnitude model [8]; an intensity PV model [11]; and a finite Gaussian mixture model where each tissue compartment is modelled as a single Gaussian distribution. In order for the models to be assessed, synthetic two tissue 3-D PV data with clinically realistic parameters was generated, with a series of concentric ellipsoid shells for the alternating tissue compartments. The PV voxels were simulated by anisotropic low pass filtering and then sub-sampling a high-resolution data space to create a lower-resolution data space, where one voxel in the low-resolution data space corresponds to 128 voxels in the high-resolution data space. Qualitative assessment was also undertaken using data from [15].

3. Discussion and Results

The results of fitting equations (7) and (8) to simulated and real data can be seen in figure 1, which illustrate the validity of the derived model and synthesized PV data. Results of applying the classifiers to the synthetic PV data can be seen in figure 2, demonstrating apparently superior performance of the 3-D model classifier for this synthetic PV data set. It is worth noting that if the Gaussian or Intensity PV model classifiers are applied to a three-tissue class problem, then intensity information alone will not be able to correctly classify PV voxels that are composed of the two intensity extremal tissue classes (such as CSF and WM on MRI T1 data), as these PV voxels will have intensities similar to the central tissue class.

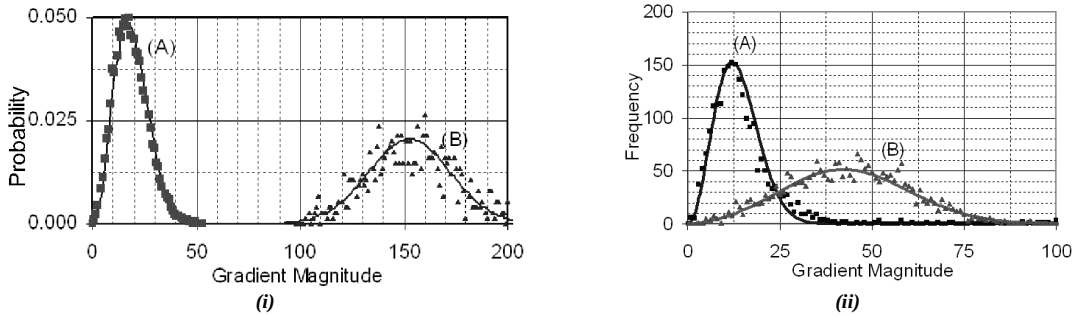


Figure 1, Comparison of derived parametric models with actual data points for 3-D gradient magnitude. (i) is for the synthetic 3-D PV data. (ii) is for clinical NMR data. Curves and points labelled (A) are for pure tissues (eq. 7) while curves and points labelled (B) correspond to PV voxels (eq. 8). Both plots illustrate good fits to both equations for real and synthetic MRI T1 gradient magnitude data

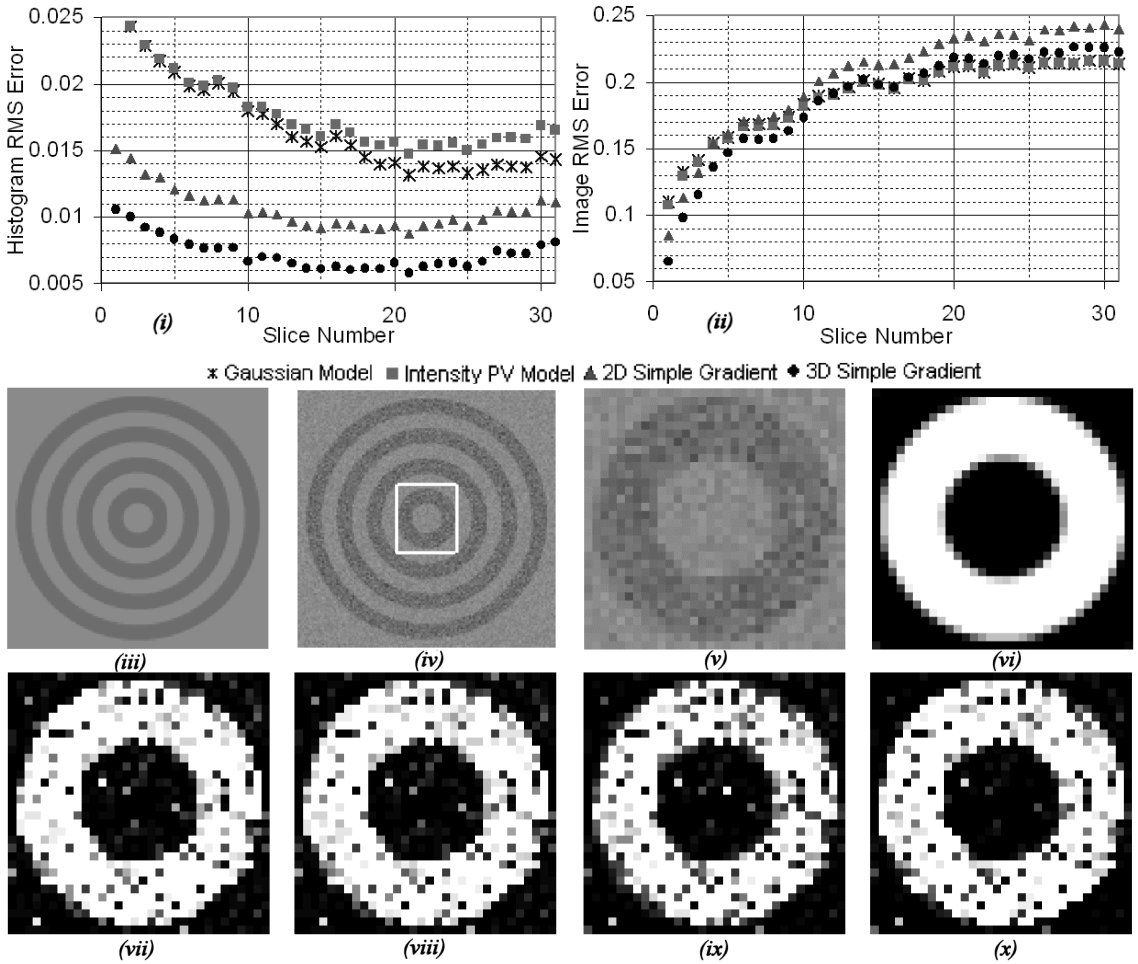


Figure 2, Performance comparison of four PV Bayesian classifiers applied to 3-D synthetic PV data (containing two tissues alternating between concentric ellipsoid shells), with parameters similar to clinical neurological MRI data parameters. Plots (i) and (ii) illustrate the performance of the classifiers for slices 1-31 of the symmetric synthetic PV data with 62 slices. (iii) is an exemplar slice through the synthetic PV sub-sampled data without noise (ground truth). (iv) is the same slice, but through the noisy data set. (v) is the inset of (iv). (vi) is the corresponding region but of the ground truth image in (iii). Images (vii) – (x) are the output of the four classifiers for the same region as shown in (v) and (vi): (vii) Gaussian finite mixture model; (viii) Intensity PV model; (ix) 2-D model; (x) 3-D model. Images (vi)-(x) have been scaled to 8 bits, to display the PV voxels. From these results, especially from (i), it appears that the 3-D model has a superior classification performance over the other models for this particular synthetic PV data set.

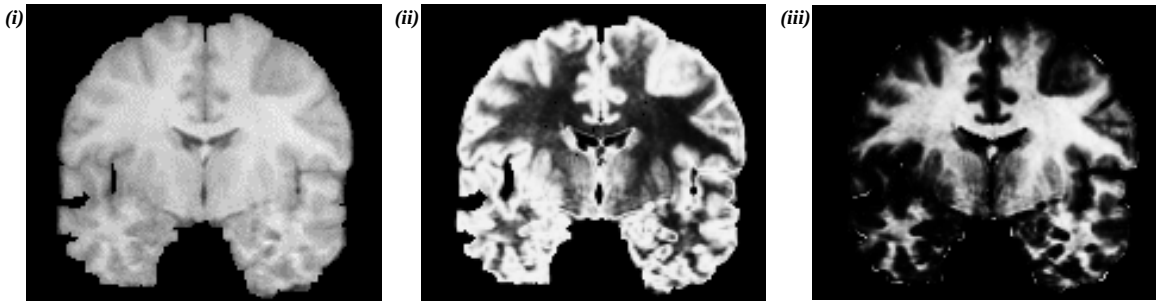


Figure 3, Classification results of T1 MRI data using the 3-D gradient magnitude model. (i) is an exemplar slice taken from a T1 MRI data scan [15]. (ii) & (iii) are the classification results for WM and GM. Skull stripping was performed using the technique in [16] and field inhomogeneities were ignored.

The 3-D model Bayes classifier was then applied to an exemplar MRI T1 dataset [15], using a three-tissue model, to classify WM, GM and CSF PV voxels. A slice from this data set and the corresponding probability maps for WM and GM can be seen in figure 3.

4. Conclusions

This work has demonstrated the potential benefit of 3-D statistical modeling of the gradient magnitude in classification of data affected by the PV effect. It also appears that data created through a 3-D image acquisition process can be classified with improved performance using statistical models that take account of the extra dimension.

Acknowledgements

John Chiverton would like to thank Hataikan Porncharoensin and Barbara Podda for useful discussions and the EPSRC for funding.

References

1. E. M. Haacke, R.W. Brown, M.R. Thompson, *et al*. "Magnetic Resonance Imaging, Physical Principles and Sequence Design". Wiley-Liss, 1999.
2. D.H. Laidlaw, K.W. Fleischer & A.H. Barr. "Partial-Volume Bayesian Classification of Material Mixtures in MR Volume Data Using Voxel Histograms". *IEEE Transactions on Medical Imaging*, 17(1), pp. 74-86, 1998.
3. M. Pokric, N.A. Thacker, M.L.J. Scott, *et al*. "The Importance of Partial Voluming in Multi-dimensional Medical Image Segmentation". In *Proceedings of MICCAI 2001*, pp. 1293-1294, Springer-Verlag, 2001.
4. M. Petrou & P. Bosdogianni. "Image Processing, The Fundamentals". Wiley, 1999.
5. H. Voorhees & T. Poggio. "Detecting Textons and Texture Boundaries in Natural Images". In *Proceedings of the International Conference on Computer Vision*, pp. 250-258, 1987.
6. E.R. Hancock & J. Kittler. "Adaptive estimation of hystersis thresholds". In *Proceedings of CVPR-91*, pp. 196-201, 1991.
7. N. Amodaj & M. Popovic. "Adaptive threshold selection for directional derivative edge detector". In *Proceedings of the 1990 Bilkent Communication, Control and Signal Processing Conference*, Ankara, Turkey, pp. 1619-1625, 1990.
8. D.C. Williamson, N.A. Thacker, S.R. Williams, *et al*. "Partial Volume Tissue Segmentation using Grey-Level Gradient". In *Proceedings of MIUA 2002*.
9. R. Bracho & A.C. Sanderson. "Segmentation of images based on intensity gradient information". In *Proceedings of CVPR-85*, pp. 341-347, 1985.
10. A. Papoulis. "Probability, Random Variables, and Stochastic Processes". McGraw-Hill, 1991.
11. E.A. Vokurka, A. Herwadkar, N.A. Thacker, *et al*. "Using Bayesian Tissue Classification to Improve the Accuracy of Vestibular Schwannoma Volume and Growth Measurement". *American Journal of Neuroradiology*, 23, pp. 459-467, March 2002.
12. J.P. Chiverton & K. Wells. "Volumetric Partial Volume Quantification via a Statistical Model of 3D Voxel Gradient Magnitude". *IEEE NSS & MIC Conference, Rome 2004.*
13. A.H. Andersen & J.E. Kirsch. "Analysis of noise in phase contrast MR imaging". *Medical Physics*, 23(6), pp. 857-869, 1996.
14. J. Zhong, A.M. Seifalian, G.E. Salerud, *et al*. "A mathematical analysis on the biological zero problem in laser Doppler flowmetry". *IEEE Transactions on Biomedical Engineering*, 45(3), pp. 354-364, 1998.
15. "The 20 normal MR brain data sets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at http://www.cma.mgh.harvard.edu/ibsr/".
16. J.P. Chiverton, B. Podda, C. Chen, *et al*. "Fully Automatic Skull Stripping of Routine Clinical Neurological NMR Data". *IEEE NSS & MIC Conference, Rome 2004.*

Adaptive Contour Segmentation for Closely Packed Cytometric Assay

W.F. Clocksin
Department of Computing, Oxford Brookes University
Wheatley, Oxford OX33 1HX, U.K.
wfc@brookes.ac.uk

Abstract. We present a segmentation method that works for overlapping and closely packed nuclei in noisy images that have high variation in background intensity. The method has been tested on fluorescence in-situ hybridisation images of interphase leucocyte nuclei. Accurate segmentation is required in support of an automatic procedure for assaying telomere content on a per area per nucleus basis. The method first finds a single seed point for each nucleus that uniquely identifies that nucleus. Seed points are located by an efficient iterative mode-finding algorithm based on robust nonparametric density estimation. Acting simultaneously on all nuclei in the image, and using the seed points as origins, flexible closed contours are dilated until each nucleus is circumscribed. Unlike previous approaches, the contour equations include a repulsive term that prevents different contours from intersecting, thereby preserving the identity of nearby or overlapping nuclei, and the contour is adaptively remeshed for greater efficiency. The advantages of this method from an implementation point of view are that the computation of seed points and contours is highly efficient and robust compared with alternative approaches. The method is illustrated using data from a clinical pilot study.

1 Introduction

Finding the boundaries of whole cells and cell nuclei – segmentation – is an essential precursor to certain types of automated cytometry. For example, whole cells need to be segmented in order to quantify the expression of cytoplasmically localised proteins [1] and to characterise cell morphology. Also, in fluorescence in-situ hybridization (FISH) [2] images, chromosome signal counts are valid only for signals contained within the nuclear boundary [3]. The boundary can be found manually, for example by drawing on the image with a mouse, but an automatic method is preferable when many nuclei must be segmented. A number of algorithms have been developed to segment nuclei based on simple thresholding [3], the watershed algorithm [4, 5], level sets [1], and curve fitting using a minimum error criterion [6].

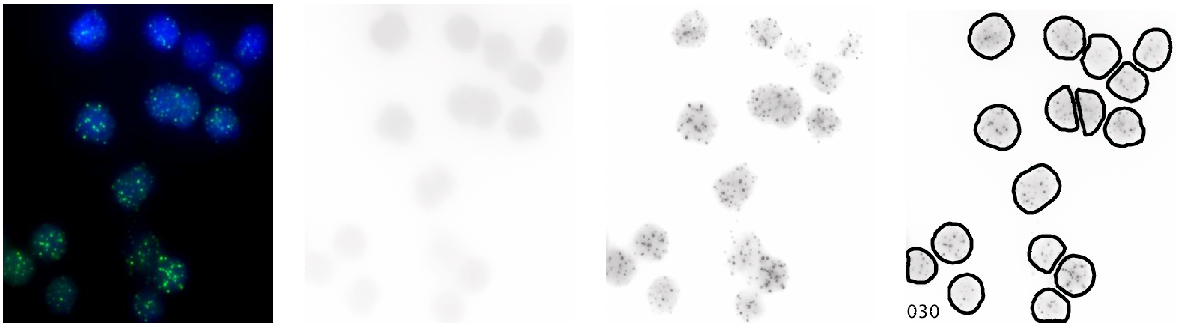
The new segmentation method introduced here was motivated by the need to process leucocyte images that routinely contained nonuniform background fluorescence and several closely packed cells (see Figure 1). A pilot application of standard methods, such as those cited above, did not yield good results because the data fail to conform to assumptions. For example, the background illumination and nucleus brightness profile varies across the image in such a way that watershed algorithms fail to work reliably, and the desire for automatic operation rules out the use of a manually assisted 'marker' watershed algorithm [7]. Applying a sequence of pre-processing operations to 'correct' the image is a possible alternative, but can lead to the introduction of additional artifacts, and in any case is an ad hoc procedure giving no promises of general applicability. Furthermore, because our images subtend a large area at high resolution, we wished to avoid the processing costs of level set methods that rely on local gradient flow. We therefore recast the nucleus finding problem in terms of probabilistic robust estimation, a relatively new approach which is beginning to find widespread application in computer vision and image processing [8].

2 Image Analysis

Analysis of the FISH images is intended for a fully automatic screening system, in which no manual intervention is necessary, and for which multiple nuclei may appear in a given image. Processing of the input RGB image treats the blue and green channels separately (the red channel is reserved for another fluorescent probe not relevant to this paper). The blue channel shows nuclei material fluorescence that can be used to localise the nucleus boundaries. The green channel shows telomere fluorescence to be assayed once nucleus boundaries are found. Although telomeres are contained within the nucleus boundary, telomere signals are not sufficiently informative to localise nucleus boundaries, particularly when nuclei are closely packed.

2.1 Seed Finding

The first step is to locate the approximate central positions of nuclei, called seeds. Seed finding is accomplished by a efficient iterative nonparametric clustering algorithm. In the images we process, nuclei contain local maxima



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

Figure 1. (a) Image of leucocyte nuclei with telomere fluorescence. (b) The blue channel, showing DAPI-stained nuclei fluorescence with nonuniform intensity. (c) The green channel, showing telomere fluorescence. (d) Nucleus boundaries found by the algorithm, superimposed on the B+G channels. The 030 refers to the number of iterations.

in the blue image channel (DAPI staining), so the image is processed to find candidate seed points as follows. The region of blue surrounding each pixel in the image is matched with a dome-shaped filter having a radius of 8 pixels. A suprathreshold filter amplitude indicates a candidate point. A seed point is found by clustering candidate points, as follows. The seed finding algorithm is based on the mean shift procedure, originally proposed in 1975 [9], and largely neglected until recently [10], when there has been a revival of interest [11]. The mean shift algorithm accomplishes the fusion of uncertain measurements arising from an unknown number of sources. It is also highly efficient, and a robust method that tolerates the presence of outliers in the data. Given a set of n d-dimensional observations $\mathbf{x}_i$ for $i = 1, 2, ..., n$, let $p(\mathbf{x})$ be the probability density function at $\mathbf{x}$. It is desired to find the modes of the density function, which in our application are associated with nuclei. Following [9], one method of finding modes is to iteratively move each observation a small step in the direction of the density gradient until tight clusters of observations result near the modes. Each observation can be moved by an amount proportional to the density gradient at the observation point. Let $\mathbf{x}_i^t$ be the ith observation at iteration t, and set $\mathbf{x}_i^0 \equiv \mathbf{x}_i$. This algorithm can be expressed as $\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + a\nabla \ln p(\mathbf{x}_i^t)$ where a is an appropriately chosen positive constant to guarantee convergence. This is the d-dimensional version of the linear iteration technique for finding the roots of the equation $\nabla_x p(x) = 0$, and equivalently the modes of the mixture density $p(\mathbf{x})$. Unfortunately the density p is not known, but it can be shown [9] that the normalized gradient of the density at $\mathbf{x}$ can be estimated from the observations by the sample mean shift: $\nabla \ln p(\mathbf{x}) \equiv \frac{d+2}{kh^2} \sum_{\mathbf{x}_i \in S_h(\mathbf{x})} (\mathbf{x}_i - \mathbf{x})$, where k is the number of of observations $\mathbf{x}_i$ in a small region $S_h(\mathbf{x})$ about $\mathbf{x}$. Region $S_h(\mathbf{x})$ includes points $\mathbf{x}_i$ for which $|\mathbf{x}_i - \mathbf{x}| \leq h^2$. Parameter h determines the amount of smoothing of the density, and correspondingly the elimination of modes that are too narrow or too close to other modes. According to [9], a good choice of parameter a is $h^2/(d+2)$, so substitution of the above and rearranging terms yields the the mean-shift clustering algorithm $\mathbf{x}_i^{t+1} = \frac{1}{k} \sum_{\mathbf{x}_j^t \in S(\mathbf{x}_i^t)} \mathbf{x}_j^t$. This algorithm moves each observation $\mathbf{x}_i$ to the sample mean of the observations within the region S around it. As soon as some observations lie within distance h of one another, the next iteration will move them to a common point, their sample mean. We have improved the efficiency of the algorithm by using a reweighting scheme, in which a weight term w_i is associated with each observation. At $t = 0$ all w_i are initialized to 1. If an iteration moves a number of observations to a common point, all but one of the observations are removed, and remaining observation is reweighted by adding to its weight the weights of the points with which it merged. For the next iterations, observations are resampled from those remaining from the previous step. The algorithm terminates when no updates can be made, in which case only isolated points remain, indicating the locations of the modes. The weight indicates the number of observations that moved to the mode, so the weight can be used as an indicator of mode confidence.

2.2 Contour Models

The next step is to find the boundary of each nucleus by expanding a closed deformable contour. Deformable contour models [12] are a powerful technique for the extraction of boundaries from images, but most existing methods have three well-known shortcomings: sensitivity to initialisation, inadaptibility to boundary concavities, and intersection of contours. These problems have led to the use of alternatives such as gradient-driven curvature flows [1], which do not have strict initialisation requirements and which have the resolution to match complex boundaries. However, the flow algorithms are based on numerical approximation of partial differential equations at each image pixel, which is computationally expensive. By contrast, the approach tried here is to address the shortcomings of contour models by (a) using a force model that is particular to the problem of finding the boundaries of nuclei, (b)

adaptively remeshing the contour to increase resolution near the target boundary, and (c) adding a repulsive force to the contour model to prevent the intersection of boundaries. Our contour model is a 2D version of a discrete deformable polyhedral model [13] in which the radii to vertices independently lengthen or shorten according to force constraints until the boundary of the object is reached. A variety of closed contours can be expressed, including concave ones, sufficient for outlining cell nuclei. A rough estimate of the object is quickly generated using a low resolution model, then the desired level of detail is achieved by remeshing.

2.3 Contour Finding

Given a set of m seed points found in the image, the contour finding algorithm begins by assigning an initial contour model C_i, $i = 0 \ldots m-1$ to each seed point. Each model C_i contains a set of n contour points $\boldsymbol{p}_{ij}$, $j = 0 \ldots n-1$, an origin $\boldsymbol{o}_i$, and a threshold θ_i. A contour point $\boldsymbol{p}_{ij}$ is conveniently represented as a radius of length r_{ij} from the origin $\boldsymbol{o}_i$. The n radii of each contour model C_i are set at equal intervals of $n/2\pi$ radians about the origin $\boldsymbol{o}_i$. With each r_{ij} set to the same small constant, the initial state of each model resembles a regular n-gon centred at a seed point. The energy functional minimised for each contour point is $E_{\text{total}} = E_{\text{internal}} + E_{\text{external}}$, where elasticity and stiffness forces are given by $E_{\text{internal}} = w_{\text{elastic}} \frac{d\boldsymbol{p}_j}{dj} + w_{\text{stiff}} \frac{d^2\boldsymbol{p}_j}{dj^2}$ with weighting parameters w_{elastic} and w_{stiff}. Some external forces require measurements of the blue channel of the image, as this channel is associated with the DAPI stained nucleus body. The intensity of the blue channel at location $\boldsymbol{v}$ is notated as $B[\boldsymbol{v}]$. In all uses of $B[]$, smoothed intensities are used, so $B[\boldsymbol{v}]$ actually gives the mean intensity in a small neighbourhood around $\boldsymbol{v}$. The external force is given by $E_{\text{external}} = E_{\text{grad}} + E_{\text{level}} + E_{\text{rep}}$ and includes three measurements: E_{grad} the local image derivative of the blue channel $B[]$, E_{level} the difference between the blue channel value at $\boldsymbol{p}_j$ and a threshold θ_i unique to the ith contour model in the image, and E_{rep} the repulsive force. The image-related forces are calculated with respect to the radius $\boldsymbol{r}_j$ of point j: $E_{\text{grad}} = w_{\text{grad}} \frac{dB[\boldsymbol{r}_j]}{dr}$, and $E_{\text{level}} = w_{\text{thresh}}(B[\boldsymbol{r}_j] - \theta_i)$. Because the blue channel has higher amplitude at the centre of the nucleus, maximising E_{grad} causes the contour to fall away from the nucleus centre and relax near the nucleus boundary. To prevent situations where $E_{\text{grad}} = 0$ causes the evolution to stick at local minima, E_{level} causes the contour to seek an absolute level θ_i. From experience with many images, a satisfactory threshold has been found to be $\theta_i = 0.2\pi B[\boldsymbol{o}_i]$, where $B[\boldsymbol{o}_i]$ is the blue value at the origin of the contour. Note that because the contour origin is updated at each iteration, this definition causes θ_i to be recalculated also. The weight w_{thresh} is given a small value so that the primary effect of E_{level} is to draw the contour away from local minima of the intensity gradient. If E_{level} is set too high, the contour may seek an unsuitable level in the presence of nearby nuclei and uneven background intensity. The repulsive force is calculated by testing whether a contour point lies within the area bounded by one of the other contour models. Each of the $m-1$ other contour models is tested. The repulsive force is defined as $E_{\text{rep}} = w_{\text{rep}} f(e_j)$ where e_j is the distance that the contour point j penetrates into the area of another contour model, and $f()$ is a nonlinear sigmoid smoothing function that limits oscillation. When all the forces are calculated for a contour point, the sum of forces is added to the radius defining the contour point. To preserve the integrity of the evolution, a two-pass update is performed. The first pass calculates the forces for all contour points, and the second pass updates all the contour point locations with the net force. When all contour point locations are updated, each contour is adaptively remeshed and its origin re-estimated.

3 Results

The segmentation algorithm has been applied to a set of fluorescence in-situ hybridisation (FISH) images of nuclei. The task is to distinguish between samples originating from individuals of different ages. Telomeres – the repeated DNA sequences that cap the ends of chromosomes – contain the same G-rich DNA core repeat. Telomeres become shorter in length as the individual ages, and depending on age, variation in repeat length of the telomeres is estimated to be on the order of 2^{11} to 2^{15} basepairs of DNA. FISH images were obtained from M. Hultén and S. Dhanjal (Univ. Warwick) who used a procedure summarised briefly as follows. Blood samples were obtained from normal adults in the age range 20-70 years, and were prepared using standard FISH techniques for the identification of telomeric material using a pantelomeric DNA probe (DAKO, Denmark). Slides were imaged using a Zeiss axioplan epifluorescence microscope with 100× objective. Filters associated with the FISH probes were used, and the images acquired using a CCD camera with SmartCapture software (Vysis, UK). Blood samples were then subjected to a telomere depletion assay (TDA) in which enzymatic digestion of telomeric DNA is carried out using Bal 31 enzyme (New England Biolabs, UK), followed by the FISH procedure and slide imaging as before. The aim of the TDA is to digest telomeric material, which should result in lower telomere measurements.

By manual inspection of all the images encountered so far, it is very difficult to distinguish between those that originate from younger subjects (expected to have longer telomeres) in relation to older subjects. To determine whether nuclei could be discriminated numerically on the basis of telomere fluorescence, the segmentation algo-

rithm was applied to a test set of four images: (a) 16 nuclei from the 70-year-old female subject before TDA; (b) five other nuclei from the same subject after TDA; (c) five nuclei from the 23-year-old male subject before the TDA was applied; (d) 13 other nuclei from the same subject after TDA. The algorithm found valid boundaries for all nuclei. Within each boundary, a telomere response measure was calculated by dividing the total amount of green channel intensity within the nucleus boundary by the area enclosed by the boundary. For each image, a normal density was fitted to the means and variances of telomere response measures for each set of nuclei. These distributions confirm expectations, as the younger samples cluster to a higher telomere content than older samples, and the TDA decreases telomere content monotonically, although apparently not linearly.

4 Conclusions

The segmentation algorithm successfully segments overlapping and closely packed nuclei on the basis of DAPI fluorescence only, even when nonuniform background and noisy profiles tend to defeat standard methods. The algorithm is efficient and capable of operating without manual intervention. Contour finding is fast and reliable, but good choice of parameters for the energy functional depends on typical sizes and shapes of objects encountered. Although a usable set of parameters has been found for nuclei shapes encountered in the FISH images, more experimentation is needed to determine usability for other domains.

The system has been tested on a situation where telomere response is required on a per-nucleus basis in support of a procedure being developed for telomere content analysis. Before a clinical application can be contemplated, it will be necessary to test the system on large numbers of nuclei, thus far not available until FISH protocols are more widely established.

Acknowledgements
This paper is a short version of [14]. I thank Maj Hultén and Seema Dhanjal of the University of Warwick for providing the TDA sample images.

References

1. C. Ortiz de Solorzano, R. Malladi, S. A. Leliévre et al. "Segmentation of nuclei and cells using membrane related protein markers." *Journal of Microscopy* **201(3)**, pp. 404–415, 2001.
2. N. P. Carter. "Fluorescence in situ hybridization – state of the art." *Bioimaging* pp. 41–51, 1996.
3. B. Lerner, W. F. Clocksin, S. Dhanjal et al. "Automatic signal classification in fluorescence in-situ hybridization images." *Cytometry* **43(2)**, pp. 87–93, 2001.
4. S. Lockett & B. Herman. "Automatic detection of clustered fluorescence-stained nuclei by digital image-based cytometry." *Cytometry* **17**, pp. 1–12, 1994.
5. N. Malpica, C. Ortiz de Solorzano, J. J. Vaquero et al. "Applying watershed algorithms to the segmentation of clustered nuclei." *Cytometry* **28**, pp. 289–297, 1997.
6. P. Bamford & B. Lovell. "Unsupervised cell nucleus segmentation with active contours." *Signal Processing* **71(2)**, 1998.
7. F. Meyer & S. Beucher. "Morphological segmentation." *Journal of Visual Communication and Image Representation* **1(1)**, pp. 21–46, 1990.
8. H. Chen & P. Meer. "Robust computer vision through kernel density estimation." In *ECCV 2002, Lecture Note in Computer Science 2350*, pp. 236–250. Springer-Verlag, 2002.
9. K. Fukunaga & L. D. Hostetler. "The estimation of the gradient of a density function, with applications in pattern recognition." *IEEE Transactions on Information Theory* **IT-21(1)**, pp. 32–40, 1975.
10. Y. Chang. "Mean shift, mode seeking, and clustering." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17(8)**, pp. 790–799, 1995.
11. D. Comaniciu & P. Meer. "Mean shift: A robust approach toward feature space analysis." *IEEE Transactions on Pattern Analysis and Machine intelligence* **24(5)**, pp. 603–619, 2002.
12. M. Kass, A. Witkin & D. Terzopoulos. "Snakes: Active contour models." *International Journal of Computer Vision* **1(4)**, pp. 321–331, 1988.
13. J.-O. Lachaud & A. Montanvert. "Deformable meshes with automated topology changes for coarse-to-fine 3D surface extraction." *Medical Image Analysis* **3(2)**, pp. 187–207, 1999.
14. W. Clocksin. "Automatic estimation of overlapping nuclei with high background variation using robust estimation and flexible contour models." In *Proc. Int. Conf. on Image Analysis and Processing*, pp. 664–669. Mantova, Italy, September 2003.

Automatic Segmentation of Liver from Computerised Tomography (CT) Images

Alun Evans*, Tryphon Lambrou, Alf Linney, Andrew Todd-Pokropek

Department of Medical Physics and Bioengineering, University College London, WC1E 6JA

Abstract

Most attempts at automatic segmentation of liver tissue to date have relied on 2D, low-level segmentation techniques, such as thresholding and mathematical morphology, to obtain the basic liver structure. The derived boundary can then be smoothed or refined using more advanced methods. Here we present results that improve greatly on this previous work by using a topology adaptive *active contour model*, or *snake*, to accurately segment the liver outline from CT images.

1 Introduction

As part of the diagnosis of liver disease, a Computerised Tomography (CT) scan is taken of the patient, which the clinician then uses to assist in determining the presence and extent of the disease. Frequently the clinician is required to hand segment the liver tissue, in order to obtain further information such as liver volume, or to quantify the extent of diseased tissue. As hand-segmentation is slow and time-consuming, an automatic segmentation tool for the liver could greatly reduce the workload for the clinician.

The automatic detection of the liver from CT scans is considered one of the harder segmentation challenges in medical image processing. The difficulties arise due to large variations of liver geometry between patients, the limited contrast between the liver and the surrounding organs, and image noise [1]. Early work on liver segmentation by Bae *et al.* [2] used simple thresholding and logic functions to obtain the outline of the liver before smoothing the boundary using B-splines. Gao *et al.* [1] extended this work by using mathematical morphology on the thresholded image to separate the liver from other organs, before refining the obtained contour with a Fourier-based deformable contour model. The limitation of both of these methods is in the initial thresholding step - it is very difficult accurately to set upper and lower threshold limits that isolate the liver effectively, without including neighbouring tissues such as the kidneys and the spleen. As a result the initial starting point for the boundary refinement step is dependent on the inaccurate thresholding step. In an effort to counter this, Shimizu *et al.* [3] use the corresponding CT values from four different input images of the same liver (each at a different stage of contrast enhancement) to obtain the rough contour. The main limitation of this technique is that four complete datasets are required for effective segmentation of one liver. Clinically this requires four different CT scans of the same patient in succession (increasing the X-ray dose dramtically), and computationally it involves four times the memory and processing power that is used when analysing a single dataset.

Recent work by Qatarneh *et al.* [4] introduced active contour models (or snakes) as a stand-alone liver segmentation tool as part of their work to construct a radiation therapy planning atlas. The limitation of this technique is again the initial placement of the contour, as it has to be defined so that it is close enough to the snake boundary so that it does not get trapped in local minima corresponding to boundaries of other tissues. In practice this is both time-consuming and awkward, especially if the 'standard' organ outline predicted by the atlas needs to be modified, by hand, for each slice.

Therefore while snakes have been effectively used in liver segmentation as a boundary refinement tool, they are still heavily dependent on the problem of initial placement. Our solution to this problem is to add an inflationary force to the basic snake, similar to the one first proposed by Cohen [5], and expanded by McInerney and Terzopoulos [6] in the form of topology adaptive snakes (T-Snakes). As the snake expands from an initial seed point, the contour reparameterises at certain iterations so that the snake elements are completely reset and replaced. In our implementation, the resolution of the grid increases as the curvature of the snake increases, thus the snake is able to force its way into sharp corners. This paper presents preliminary results from 4 separate 3D patient datasets.

* Author Email: aevans@medphys.ucl.ac.uk

2 Methods

The active contour model we have developed is a modified version of the T-Snake [6], where the internal energy forces act as a smoothness constraint while an inflation force is used to push the contour towards edges within the image. The key difference between a topology adaptive model and a conventional snake is that the set of snake nodes in a T-snake does not remain constant. As the contour moves under the internal and external forces, it is *reparameterised*, at regular intervals, to a grid superimposed upon the image. At each reparameterisation step, the previous set of nodes is removed and a new node added at each point where the contour intersects with the superimposed grid. This reparameterisation overcomes aliasing problems that naturally occur with inflationary contours, and allows the contour to 'flow' into the complex shape of the liver.

There are two major benefits of using an inflationary snake to segment the liver. The first is that the interior of a healthy liver is typically more uniform than the exterior, thus there are less noisy edges that can trap nodes as they move outwards towards the edge of the liver. The second is that the snake can be initialised at almost any point within the liver, without greatly affecting the final segmentation result. The importance of the these two points should not be underestimated as they completely avoid the major stumbling block of previous liver segmentation algorithms, that of correct contour initialisation, and thus provide a faster and more robust segmentation. Our snake is defined in a similar way to the T-Snake [6], as a set of M nodes, indexed by $0,\ldots,M-1$. Associated with these nodes are time varying positions $\boldsymbol{x}_i(t) = [x_i(t), y_i(t)]$. The movement of the snake is determined by the balance of internal forces versus external forces (generated by the image data) and inflationary forces [6,7]

$$a\boldsymbol{\alpha}_i + b\boldsymbol{\beta}_i = \boldsymbol{\rho}_i + \boldsymbol{f}_i \tag{1}$$

$\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ represent the elastic and bending internal forces of the snake, the strength of which are controlled by a and b. In our implementation the external energy function $\boldsymbol{f}$ is the standard gaussian gradient function, which is greater at significant edges in the image, however we pre-process the image using a Kirsch edge-detection filter to maximise the impact of these edges. A modification to the basic snake equations is the addition of an inflationary force $\boldsymbol{\rho}$ to push the contour towards image edges

$$\boldsymbol{\rho} = qF(m(\boldsymbol{x}_i(t))\boldsymbol{n}_i(t) \tag{2}$$

where $\boldsymbol{n}_i$ is the unit normal vector to the contour at node m, and q is the amplitude of the force. The binary function $F(m(x, y))$ is based upon image intensity data and is slightly modified from the similar T-Snake function [6] in that it has two threshold levels, an upper and a lower threshold. If $m(x, y)$ is within the threshold levels, $F(m(x, y)) = 1$; otherwise it is set to -1 and the normal force is reversed. This prevents the snake from leaking into other organs in the abdomen at locations where the external image energy is not sufficient enough to stop the snake. To prevent the normal force from oscillating indefinitely between areas of intensity within/outside the threshold levels, as soon as a node begins oscillating the q amplitude value is lowered progressively towards zero.

The reparameterisation of the snake is slightly different from that presented in [6]. The grid used for reparameterisation is set to a rectangular grid for simplicity, and a major modification is that the resolution of this grid changes depending on the curvature of the snake at each individual node. The data structure in fact consists of three separate grids of decreasing cell size (increasing resolution), although number of different resolution grids could be increased if required. Depending on the curvature of the snake (calculated simply by analysing the angles between the lines connecting the nodes), the contour is reparameterised on a specific grid. If the curvature is high, the contour is reparameterised on a smaller grid size (down to 2x2 pixels); if the contour is relatively flat it is reparameterised to a larger grid size (up to 8x8 pixels). The major advantage of this novel technique is that the resolution of the snake increases at complex and highly irregular areas of the shape to be segmented, thus enabling the inflating contour to push itself into sharp corners and avoid aliasing effects that might otherwise cause a false segmentation result. In areas where the contour is relatively straight a larger grid size is used for reparameterisation and less points are required to model the shape, reducing the number of unnecessary calculations and speeding up the performance of the snake. Accurate segmentation of the liver is highly dependent on the parameters used for the snake. In all there are six parameters to be set: the internal elasticity energy, the internal bending energy, the external (image) energy, the strength of the normal force, and the upper and lower thresholds. The levels of these parameters were determined empirically to obtain the best results and to prevent the snake from leaking into adjacent organs. We found that once the optimal levels for the parameters were discovered, they required little or no change for every liver slice, even for different livers (the exception to this being particularly noisy datasets). Segmentation of a complete liver is initiated by the user selecting a seed point within the liver i.e. any slice in the volume. The snake then inflates from this seed point to segment the liver from that particular slice. From the resulting region enclosed by the snake contour, the co-

ordinates representing the centre of gravity are deduced and used as snake seed points for the slices above and below. In many scans, a lobe of the liver appears as a totally separate structure. In these cases the snake can be initialised on the lobe in much the same way as for the main body of the liver, and results added to the main segmentation.

3 Results

Complete scans from four different patients, giving a total of 501 separate liver slices, were used to test the proposed segmentation method. All the images used were 512x512 pixels and reduced 256-level greyscale. The accuracy of the segmentation can be measured by comparing to the hand-segmented data in two ways; by comparing the area enclosed by the segmented contours, and by calculating a rout-mean-square (RMS) error for the Euclidean distance between the automatic segmentation contour and the nearest point of the hand segmented contour. To compare areas, a paired t-test was used [8], and the null hypothesis in each case was "There no difference in the areas of the snake segmented liver and the hand segmented liver.", and the significant probability level (p) was set to 0.05.

Patient Number	Slices	T-value (experimental)	Critical T-value ($p < 0.05$)	RMS Error (pixels)
1	180	0.406	±1.973	4.82
2	149	-0.556	±1.976	7.57
3	122	-1.591	±1.990	5.83
4	50	-1.347	±2.009	10.22

Table 1. Results of paired t-test comparing segmented areas;

Table 1 shows the results of the t-tests and RMS errors. For each dataset, the t-value obtained is less than the t-statistic presented in the data tables. This means that we cannot reject the null hypothesis at the 0.05 level of significance, and must therefore assume that there is no detectable change, at this level of significance, in the values of the snake segmented and hand segmented data. The low RMS errors are comparable with those expected from hand segmentation. To demonstrate the improved segmentation ability provided by increasing the resolution of the grid at areas of high curvature, the liver from patient number one was automatically segmented again, using an inflationary snake that was reparameterised to a static 8x8 pixel grid i.e. the resolution of the grid was *not* changed at areas of high curvature. Table 2 shows the results of this segmentation, compared with the standard results presented in table 1.

Grid resolution	Experimental t-value	Average RMS error
Increased at areas of high curvature	0.406	4.82
Constant	-3.747	5.23

Table 2. Comparing the effect of grid resolutions on the accuracy of the segmentation.

The results in table 2 show that the magnitude of the t-value result for a constant grid resolution is larger than the critical t-value of 1.973 (shown in table 1) and thus according to this test there is a *significant difference* between the automatically segmented and hand segmented areas. This value is much larger than the t-value of 0.405, obtained from the same dataset, using a flexible grid resolution for reparameterisation. However, while there is an increase in the average RMS error for the constant grid segmentation, it is proportionally lower than the increase in the difference in areas. The reason for these differences can be seen by studying the two images in Figure 1. Figure 1(a) shows a liver slice that has been segmented using a snake with varying grid resolution. Figure 1(b) shows a liver slice that has been segmented using the snake with constant grid resolution. In both of these images the snake was initialised within the central bulk of the liver. One immediately observed difference between the images is that the snake in figure 1(b) has not inflated into the lobe at the top right hand side of the liver. The likely reason for this is that the relatively low resolution of the grid (8 pixels) has not enabled points to cluster around the narrow area of tissue connecting the lobe to the main bulk of the liver (in this slice). As a result the inflationary force is not sufficient to overcome the internal forces of the contour and the snake does pass through the narrow constriction. In figure 1(a) the ability of the contour to reparameterise to a smaller grid size (min size 2 pixels) has enabled points to cluster around areas of high curvature, and force the snake through the constriction and allow it to inflate into the lobe area. Figure 1(c) shows the ability of the increased resolution grid to more accurately segment fine structures.

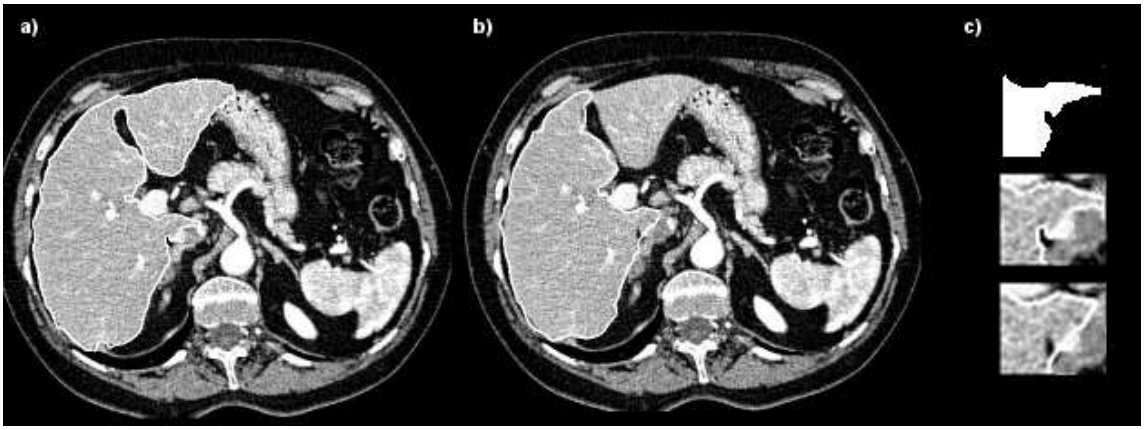


Figure 1: Two images showing the same liver slice segmented by different snakes. (a) is segmented using a varying resolution reparameterising grid, (b) is segmented using a constant grid size. c) compares the hand segmented area (top) against a close up of the same areas from image a) (middle) and image b) (bottom).

4 Conclusion

This paper presents a new method for the automatic segmentation of the liver from CT scans. It avoids the main problem affecting previous segmentation techniques, that of initialising the snake in an efficient manner, by employing an inflationary snake which reparameterises at certain iterations of the snake movement. The snake algorithm itself is a modified version of the T-Snake presented by McInerney and Terzopoulous [6] with the important added facility of increasing the resolution of the reparameterisation grid where the snake contour is highly curved. It has been demonstrated that this ability enables more accurate segmentation, as it enables the contour to extend into corners that other 'inflationary' snakes miss. While our preliminary segmentation results are very encouraging, there is still much more to be done to improve the technique. An immediate improvement is to enable the software to set snake parameters automatically for each liver slice – at the moment the parameters must be set by hand. While livers have also been segmented using orthogonal datasets (2.5D), the major goal of the project is to develop an active surface model in 3D, incorporating our existing modifications and improvements to the 2D method. This would further improve the accuracy of the segmentation of the liver, and form a firm basis for further work concerning abnormal livers.

Acknowledgements

This work was supported by EPSRC and MRC under the Interdisciplinary Research Consortium scheme - "From Medical Images and Signals to Clinical Information.

References

1. L. Gao, D. Heath, D. Kuszyk et. al. "Automatic Liver Segmentation Technique for Three-dimensional Visualisation of CT Data", *Radiology* **201**, pp 359-364, 1996.
2. K. Bae, M. Giger, C. Chen et. al. "Automatic segmentation of liver structure in CT images", *Medical Physics* **20**(1), pp 71-78, 1993.
3. A. Shimizu, T. Hitosugi, J. Nakagawa et. al. "Development of computer-aided diagnosis system for 3D multi-detector row CT images of livers", *International Congress Series* **1256**, pp 1055-1062, 2003.
4. S. Qatarneh, M. Noz, S. Hyödynmaa et. al. "Evaluation of a segmentation procedure to delineate organs for use in construction of a radiation therpay planning atlas", *International Journal of Medical Informatic,* **69**, pp 39-55, 2003.
5. L. Cohen, "On Active Contour Models and Balloons", *Computer Vision, Graphics, and Image Processing: Image Understanding* **53** (2), pp 211-218, 1991.
6. T. McInerney, & D. Terzopoulos "T-snakes: Topology adaptive snakes", *Medical Image Analysis* **4**, pp 73-91, 2000.
7. M. Kass, A. Witkin, & D. Terzopoulos "Snakes: Active Contour Models", *International Journal of Computer Vision*, **4**(1), pp 321-331, 1988.
8. R. F. Mould, *Introductory Medical Statistics.* Adam Hilger, Bristol, England, 1989.

Segmentation of Volumetric Prostate MRI Data Using Hybrid 2D+3D Shape Modeling

Yanong Zhu[a]*, Stuart Williams[b] and Reyer Zwiggelaar[c]†

[a]School of Computing Sciences, University of East Anglia, Norwich, UK
[b]Department of Radiology, Norfolk and Norwich University Hospital, Norwich, UK
[c]Department of Computer Science, University of Wales, Aberystwyth, UK

Abstract.
As a straight forward extension of 2D Active Shape Modelling (ASM), 3D ASM is capable of detecting the surface of the target object from true volumetric data. However, when the information in one dimension is sparse, the pure 3D ASM tends to be less accurate. We present a hybrid 2D+3D methodology which can deal with sparse 3D data. Based on the "locally optimal" segmentation on separate slices obtained by 2D ASM, a "global optimal" segmentation of the 3D object embedded in the data set is achieved by applying 3D ASM. Experimental results indicate that the developed approach shows equivalent precision on separate slices but higher consistency for whole volumes when compared to 2D ASM, whilst the results for whole volumes are improved when compared to the pure 3D ASM approach.

1 Introduction

Segmentation of target objects from multi-slice, i.e. 3D, data sets is of great interest. Since it is difficult to extract the 3D object directly from the data sets, slices in these data sets tend to be segmented as separated 2D images. In the past decade, Active Shape Models (ASM) [1], or generally so-called statistical shape models [2], have been widely used and shown to be a very powerful tool in 2D medical image segmentation. Despite the successful cases of applying ASM, it has the same disadvantage as other 2D segmentation approaches, i.e. they are purely 2D segmentation techniques which process the slices in a data set separately, in most cases ignoring the correlation between slices. The result might be 2-dimensionally optimal on every slice but not the ideal boundary for the whole 3-dimensional object embedded in the data set.

As a significant extension of Active Shape Models (ASM), 3D ASM has been widely investigated and discussed. Many studies build statistical models from 3D training data sets [3–7]. Most of these developments have focused on the landmark generation and model construction, with an assumption that if an efficient model has been built the use of such a model will be a relative easy task. This is true for *dense* 3D data sets in which image information in arbitrary directions can be extracted by interpolation. In practice, however, when the distance between slices is large compared to the in-slice resolution, interpolation becomes less appropriate. In this case, the 3D shapes needed to build the models can be constructed from the 2D contours but the image structure modelling and 3D ASM search become difficult.

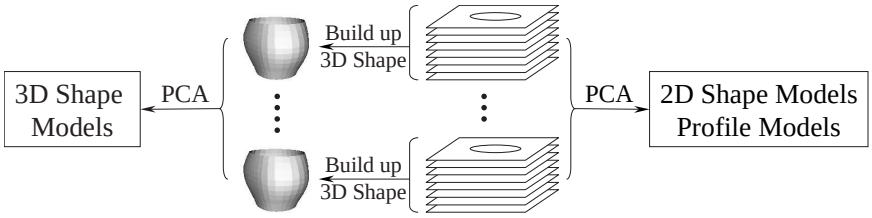
There is not much work about using of 3D ASM, or related approaches, for the segmentation of image sets. Hill et al. presented a statistical model-based technique for building 3D deformable shape templates and use these models to segment brain structures from 3D medical images [8]. Ruff et al. described a method to estimate the volume of liver from sparse planar images using deformable models [9]. Although these methods are application specific, they demonstrate the potential of 3D ASM when applied to multi-slice data sets.

2 Method

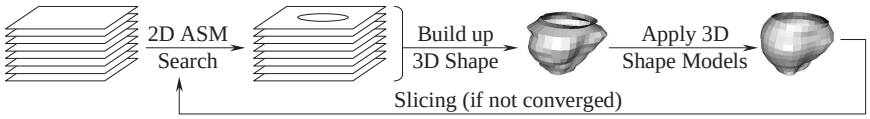
In this work, we present a new approach, based on 3D ASM, to the automatic segmentation of sparse multi-slice data sets. A 3D Statistical Shape Model (SSM) is build from the training samples to represent the shape variation of 3D object surfaces. The image structure is modelled using grey-level profiles, in the same way as 2D ASM. The 3D ASM search is driven by 2D ASM search on each slice based on the profile models built from 2D slices. Such a process ensures that 2D boundaries, which are locally correct on 2D slices, can build a valid 3D shape. In the remainder of this paper, we refer to the developed methodology as 3D ASM or 2D+3D hybrid ASM, whilst the original 3D ASM is refered to as *pure* 3D ASM.

*Email: yz@cmp.uea.ac.uk
†Email: rrz@aber.ac.uk
This project is funded by Prostate Research Campaign UK.

An overview of model construction and image segmentation is given in Fig. 1. Details are provided in Secs 2.1 and 2.2.



(a) Building 2D and 3D models



(b) Segmentation using the 2D+3D hybrid model

Figure 1. Summary of model construction and image segmentation.

2.1 3D Shape Model Construction

Each sample in the training set for 3D statistical shape modelling is the 3D object surface in a data set. To build this 3D surface, each slice that contains the target object is manually annotated and the object boundary is represented by a contour $C_i = \{p_{ij}; j \in [1, N_l]\}$ for all $i \in [1, N_c]$, where p_{ij} is the j^{th} landmark on the i^{th} contour, N_l is the number of landmarks on each slice and N_c is the number of slices in the data set. A triangulation algorithm [10] is used to incorporate 2D contours into a 3D shape in the form of a triangle mesh $M = \{v_{ij}\}$ where $v_{ij} = p_{ij}$ are the vertices of the mesh.

The 3D shapes are normalized using Principal Axes to ensure they occupy a common space, i.e. equivalent landmarks on different surfaces are at corresponding locations. We use Principal Component Analysis (PCA) to find the principal axes of a 3D shape. The point distribution of the center of gravity of each planar contour provides the first principal axis. Subsequently, all aligned contours are projected onto the same plane orthogonal to the first axis, to form a 2D point distribution. Of this distribution the first two mutually orthogonal principal axes form the remaining principal axes. The new coordinates system defines the uniform space, which can be used for shape normalization and alignment.

After the training shapes are transformed to the common space, uniform sampling is performed to generate a landmark mesh of size $N_{vert} \times N_{hor}$ for each training shape. The sampling frequency (or the mesh size) needs to be high enough to match the vertex density and preserve the quality. These 3D landmarks in the uniform space are then transformed back to the original space of the training sample. Thus a shape $\mathbf{u}$ in the training set can be treated as a $3 \times N_{vert} \times N_{hor}$ dimensional vector. The training shapes need to be aligned to minimize the effects of different pose and size. To align two 3D shapes $\mathbf{u}$ and $\mathbf{u}'$, we try to find a linear transformation $\mathbf{u}' = T(\mathbf{u})$ which minimizes the distance between the two shapes. The transformation is defined as the combination of a rotation $\mathbf{R}$, a scaling $\mathbf{s}$ and a translation (x_t, y_t, z_t). The principal axes are used to rotate the shapes into common coordinates. Let $\mathbf{v}_x$, $\mathbf{v}_y$ and $\mathbf{v}_z$ be the unit vectors of the principal axes of shape $\mathbf{u}$, thus $\mathbf{V} = ([\mathbf{v}_x\ \mathbf{v}_y\ \mathbf{v}_z]^T)^{-1}$ defines the rotation from original space to normalized space, so the rotation from $\mathbf{u}$ to $\mathbf{u}'$ is $\mathbf{R} = \mathbf{V} \cdot \mathbf{V}'^{-1}$, where $\mathbf{V}'^{-1}$ is the rotation of $\mathbf{u}'$ from its local space to the normalized space. The rotated shape $\mathbf{u}$ is then scaled so that its volume is equal to $\mathbf{u}'$. Let $|\mathbf{u}|$ be the volume of a 3D shape $\mathbf{u}$, the scaling factor $\mathbf{s}$ is given by $\mathbf{s} = \sqrt[3]{|\mathbf{u}'|/|\mathbf{u}|}$. Finally, the centre of gravity of $\mathbf{u}$ and $\mathbf{u}'$ are calculated, denoted $(\dot{x}, \dot{y}, \dot{z})$ and $(\dot{x}', \dot{y}', \dot{z}')$. The translation is given by $(x_t, y_t, z_t) = (\dot{x}' - \dot{x}, \dot{y}' - \dot{y}, \dot{z}' - \dot{z})$. The alignment of the whole training set is performed by an iterative method which minimizes a weighted sum of squares of distances between equivalent landmarks on different shapes [1].

After applying PCA to the aligned 3D training samples, we can obtain a 3D statistical shape model. An instance of the model $\mathbf{u}$ can be generated by:

$$\mathbf{u} = \overline{\mathbf{u}} + \hat{\mathbf{P}}_s \hat{\mathbf{b}}_s \tag{1}$$

where $\overline{\mathbf{u}}$ is the mean shape, $\hat{\mathbf{P}}_s$ is a set of orthogonal modes of shape variation, and $\hat{\mathbf{b}}_s$ is a vector of shape parameters. Varying the elements of $\hat{\mathbf{b}}_s$ moves the landmarks within the distribution and generates shape instances from the model.

2.2 Segmentation

We concentrate on the segmentation of multi-slice data sets in which the gap between adjacent slices is significantly larger than the in-slice resolution. The lack of information between slices limits the direction and position to extract intensity profiles. As a result, pure 3D ASM search approaches, in which grey-level values are extracted along the profiles across each 3D landmark and *normal* to the 3D surface, are not suitable for this type of data.

In the developed method, the segmentation process is performed in an iterative way, as demonstrated in Fig. 1 (b). A complete 2D ASM search on all slices is considered as the first step in one iteration of 3D ASM search. When a 3D shape instance is placed on the target data set, the initial 3D shape is first translated and scaled so that it is just within the range of these slices, to make sure that all the slices that contain the target object are covered. The intersecting contours of this instance and the data slices are used as the initial 2D shape instances for all slices. Then 2D ASM search is performed to find the best position for each contour. A 3D shape is constructed from these contours and then normalized using the method presented in Sec. 2.1. New landmarks are extracted from the normalized shape and subsequently mapped back onto the 3D shape. The result is an updated 3D shape instance. A set of parameters $\hat{\mathbf{b}}_s$ as defined in Eq. (1) are estimated to best match a model instance to the new found shape. An iterative approach to parameter estimation is used [1].

3 Experiments and Results

Experiments were carried out to segment the prostate gland from sequential Magnetic Resonance Images (MRI). The data consists of 19 prostate MRI volumes. Field of view 24×24 cm, matrix 256×512, slice thickness 3 mm with an inter-slice gap of 0.5 mm. Hence the in-slice resolution is 0.46875×0.46875 mm, while the inter-slice resolution is 3.5 mm. Different types of prostate abnormalities are included. All images were manually annotated by an expert radiologist. Since the central gland is nearly oval-shaped, we choose the four intersection points of the outline and the vertical and horizontal axes through the centroid as the key landmarks. On each of the four outline sections, seven landmarks were evenly distributed, so contours are represented by 32 landmarks.

For 3D shape modelling, the 3D surfaces of the prostate in all data sets were constructed. Since the prostate gland is almost symmetric, the mirror of a 3D shape in the plane $x = 0$ can be treated as a new instance and is also used in training. Hence in total we have 38 samples in the training set. The 3D Shape Model is built from these shapes using PDM. Such a model is capable of representing the shape variation of the prostate gland.

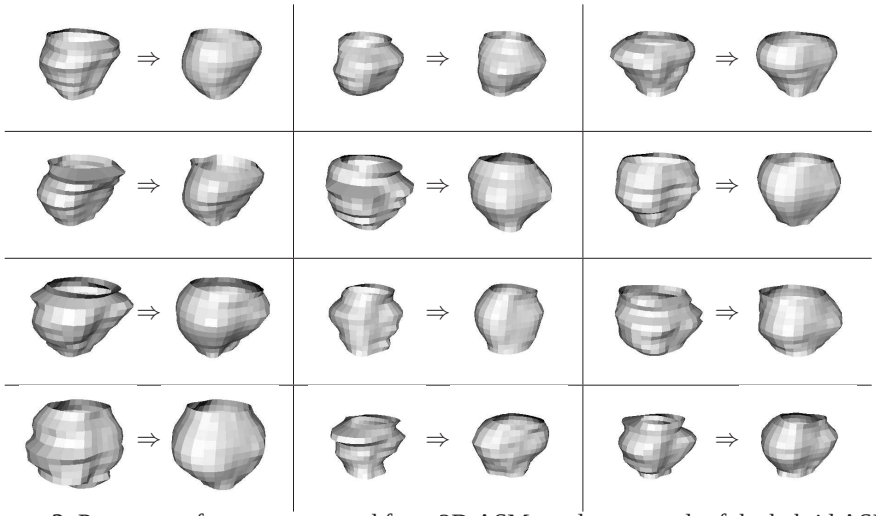


Figure 2. Prostate surface reconstructed from 2D ASM results vs. result of the hybrid ASM.

The hybrid 2D+3D model was applied to detect the prostate surface from MR data sets. Leave-one-patient-out experiments were carried out to investigate the performance of the developed approach. A comparison of segmen-

tation results between 2D and the developed method in twelve randomly selected cases is shown in Fig. 2. On the left of each pair of shapes is the initial 3D surface constructed from the 2D ASM results in the first iteration. The constructed surface does not represent a valid prostate shape, though the 2D ASM results might be optimal on each slice. On the right is the 3D shape after the last iteration. This shape shows high similarity to the left one and is a valid prostate shape since it is generated from the training set using the 3D shape model.

We use the Root Mean-Square Distance (RMSD) to evaluate the developed approach. Using manual annotations as the gold standard, we compare the results with 2D ASM and pure 3D ASM segmentations. To make a statistical evaluation, the mean and standard deviation RMSD for all landmarks are calculated and presented in Table 1. The hybrid 2D+3D ASM shows the lowest values for both the mean and standard deviation RMSD, indicating the highest precision and robustness.

Method	RMSD Mean	RMSD SD
2D ASM	6.1830	3.5326
Pure 3D ASM	13.0643	5.8128
Hybrid ASM	5.4497	2.9289

Table 1. Mean and Standard Deviation (SD) RMSD.

4 Conclusions and Discussion

We have demonstrated a hybrid 2D+3D ASM-based methodology to segment object of interest from sparse multi-slice data sets. The segmentation is driven by 2D ASM search on each slice and the valid surface is generated from the 3D shape model by updating model parameters. The results indicate that the developed methodology shows equivalent precision of segmentation on individual slices and performs more consistently with regard to capturing the 3D surface of the target object when compared to 2D ASM. A comparison with pure 3D ASM results indicates the improved performance of the developed methodology when applied to sparse volumetric data.

The method can be improved in several aspects. As discussed in other work, landmarking is a key point to improve the compactness and simplicity of shape models. On the other hand, both 2D ASM and the developed method, as shown in our experiments, tend to over-estimate the prostate gland. This might be caused by either the 2D ASM during the search stage or the parameter estimation in the updating stage. The method of aligning two 3D shapes can be evaluated and improved to minimize the error.

In spite of the improved performance, this method has its limitations. In our experiments of prostate segmentation, the prostate on transectional slices is almost oval-shaped. This feature ensures that the method we have proposed in Sec. 2.1 to calculate the principal axes is valid. In other cases, a more generalized alignment method should be applied. This is one of the most important aspects in our future work.

References

1. T. F. Cootes, A. Hill, C. J. Taylor et al. "The use of active shape models for locating structures in medical images." *Image and Vision Computing* **12(6)**, pp. 355–366, 1994.
2. I. Dryden & K. V. Mardia. *The Statistical Analysis of Shape*. Wiley, London, 1998.
3. R. H. Davies, C. J. Twining, T. F. Cootes et al. "3D statistical shape models using direct optimisation of description length." *Lecture Notes in Computer Science* **2353(3)**, pp. 3–21, 2002.
4. A. D. Brett & C. J. Taylor. "A method of automated landmark generation for automated 3D PDM construction." *Image and Vision Computing* **18(9)**, pp. 739–748, 2000.
5. C. Lorenz & N. Krahnstöever. "Generation of point-based 3D statistical shape models for anatomical objects." *Computer Vision and Image Understanding* **77(2)**, pp. 175–191, 2000.
6. M. Styner, K. Rajamani, L. P. Nolte et al. "Evaluation of 3D correspondence methods for model building." *Lecture Notes in Computer Science* **2732**, pp. 63–75, 2003.
7. A. F. Frangi, D. Rueckert, J. A. Schnabel et al. "Automatic 3D ASM construction via atlas-based landmarking and volumetric elastic registration." *Lecture Notes in Computer Science* **2082**, pp. 78–91, 2001.
8. A. Hill, A. Thornham & C. J. Taylor. "Model-based interpretation of 3D medical images." *Proceedings of British Machine Vision Conference* pp. 339–348, 1992.
9. C. F. Ruff, S. W. Hughes & D. J. Hawkes. "Volume estimation for sparse planar images using deformable models." *Proceedings of British Machine Vision Conference* pp. 90–99, 1997.
10. M. J. Zyada, A. R. Jones & P. G. Hogan. "Surface construction from planar contours." *Computing & Graphics* **11**, pp. 393–408, 1987.

Automatic segmentation of T_1 parametric maps of breast MR images via a hidden Markov random field

Georgios Ketsetzis[a*] and Michael Brady[a]

[a]Medical Vision Laboratory, Department of Engineering Science, Robotics Research Group, Ewert House, Ewert Place, Summertown, Oxford OX2 7BZ

Abstract

In this paper we apply Hidden Markov Random Fields to tissue segmentation of T_1 parametric maps of Magnetic Resonance (MR) breast images. The proposed algorithm incorporates an initial correction of the bias field, and automatic background removal. Maximum likelihood is used to provide an initial segmentation/classification. This classification allows for tissue parameter estimation, providing an initialization of probabilistic moments that are incorporated into a Gaussian probability model for each tissue class. The class labels follow a Gibbs distribution and the energy function is a sum of potentials taken from a multilevel logistic model for Markov Random Fields. The segmentation is obtained via maximization of the posterior probability distribution function and the solution is found by application of the Simulated Annealing (SA) algorithm. After each SA iteration, the tissue parameters are updated. The process continues iteratively until convergence. The segmentation results demonstrate anatomically plausible segmentation of T_1 parametric maps of breast MR images and we expect the method to aid real time automatic segmentation of breast tissue, particularly in diagnosis of pathology.

1 Introduction

Segmentation of breast Magnetic Resonance (MR) images into different tissue classes, such as fat, healthy and malignant tissues is an important task during the diagnostic process. Breast MR images have a number of features. They are statistically complex as they are not piecewise constant and they possess a large number of classes. Moreover, MR breast images do not have high contrast between different tissues. By carefully choosing pulse sequence parameters and gradients, it is possible to highlight different components in the object being imaged and produce high-contrast images in order to facilitate segmentation and classification [1]. On the other hand, ideal imaging conditions are never realized in practice. Electronic noise and the bias field (intensity inhomogeneities in the RF field) corrupt the quality of the image and prevent accurate segmentation.

Whilst the compensation for noise induced artifacts has already been addressed during T_1 mapping [1], the limitations caused by the bias field are overcome using the method of Styner et al. [2] that leads to automatic bias field removal. That enables reliable T_1 mapping which is required for the eventual T_1 map segmentation.

A wide variety of approaches have been proposed for image segmentation and they may be roughly divided into two categories: structural and statistical. Structural methods include various edge detection algorithms that have been applied to extract boundaries between different tissues [3]. However, such algorithms are vulnerable to artifacts and noise, and fail in breast imaging because of the anatomical tissue complexity. Region growing [4] is another popular structural approach, in which an image is divided into small regions, which can be considered as "seeds" that grow under certain criteria; but it is not considered robust. Statistical methods approach segmentation/classification from a completely different perspective. Such methods label pixels according to probability values, usually determined by the intensity distribution of the image. Thresholding methods are simple statistical methods that are unlikely to produce reliable results for breast MR images since they are not robust under the presence of noise and have no way of incorporating discrepancies resulting from PVEs. Depending on whether a specific functional form for the density model is assumed or not, a statistical approach can either be parametric or non-parametric. Maximum *a posteriori* probability (MAP) or maximum likelihood (ML) are the most usually used principles for such attempts. In such methods the probability density functions (pdfs) of the different tissue classes need to be chosen very carefully.

In contrast to conventional approaches for image segmentation, we perform tissue segmentation into a set number of classes by applying Hidden Markov Random Field - Maximum *a posteriori* (HMRF-MAP) labelling to parametric maps of T_1 images. We seek the labelling of an image that is considered to be a realization of a MRF through minimizing the risk of misclassification, which is equivalent to maximizing the posterior probability. The minimal risk estimate is known as the MAP estimate. Under the MRF model the MAP estimate may be found by minimizing the posterior energy function [5]. This minimization is achieved by applying the SA algorithm. SA converges to the global minimum and has been shown to achieve that reliably even for the most difficult minimization problems. This is the reason for its selection for the current problem, since the only alternative, the Iterated Conditional Modes (ICM) algorithm converges to a local minimum and does not therefore guarantee a final reliable segmentation in

*Corresponding Author's email: geok@robots.ox.ac.uk

such a difficult application such as the segmentation of parametric T_1 maps of breast MR images. Computational power compensates for the computational complexity of the SA and the method can be used in real time applications. The paper is organized in the following way. Section 2 provides background on HMRFs and image segmentation, followed by formulation of the problem for T_1 parametric maps of breast MR images. Section 3 provides the complete algorithm for the proposed setup and the segmentation. Finally, Section 4 demonstrates results from application of the method to real datasets and discusses the performance of the proposed method as well as future research directions.

2 Segmentation using parametric and non-parametric MRFs

The objective of our work is segmentation/classification of T_1 parametric maps of MR breast images into different types of tissues. During such a task it is important to consider pixel neighborhood information, as T_1 parametric maps arise from breast MR images that are considered to be piecewise continuous. This is achieved by modelling the tissue class distribution as a Markov Random Field. A MRF is a collection of random variables which are defined on a finite lattice, and where each variable interacts with some subgroup of that lattice termed its neighborhood. In modelling the interaction between data and model, certain parameters are required. Depending upon whether these are known or not, two paradigms result. MRF is the paradigm corresponding to *a priori* known statistical parameters. HMRF is the paradigm where statistical parameters are not known and need to be estimated. A HMRF is a stochastic process generated by a MRF whose state sequence cannot be observed directly but only through a field of observations. By imposing contextual constraints, we expect neighboring pixels to have the same class labels (in the case of piecewise constant images) or similar intensities (in the case of piecewise continuous images). This is achieved through characterizing mutual influences among pixels using conditional MRF distributions [5].

The procedure of estimating the unknown parameters for the HMRFs is known as model fitting. As we have introduced the idea of segmenting an image using HMRFs we need to introduce and make the assumption of a multivariable Gaussian emission function [5]. To fit the GHMRF at each iteration, we need an estimate of the means and covariances of the data classes. For computational efficiency, we evaluate the means and covariances using their general statistical definition. Segmentation/classification of the image is the result of MAP estimation of the tissue labels by application of the SA. We show that by incorporating, a parametric GHMRF , an initial class estimation using maximum likelihood and iterative MAP estimation using SA and mean and covariance updates, an accurate and robust segmentation of T_1 parametric maps of MR breast images is achieved.

3 Methodology

The first step in any MR breast image segmentation consists of bias field and background pixel removal. The bias field correction follows the scheme proposed in [2], where intensity inhomogeneities are assumed to form a continuous field in the image and are approximated by Legendre polynomials. The associated parameters are computed via the solution of a non-linear energy minimization problem, computed from the distance of pixel values and class means. We thus remove the frequency artifact due to inhomogeneities in the radio-frequency field. Automatic background removal is achieved by Otsu's thresholding method. Background intensities have a much lower value than breast intensities and the global threshold found by Otsu's method minimizes the intra-class variance. Automatic background removal is performed to speed up further processing of the images.

In what follows, we establish the segmentation of the T_1 parametric map using GHMRFs and SA and we adopt the same notation as in [6]. Let X be a realization of a random field $\mathbf{X}$ defined on a 2-D lattice $\mathcal{S}$, where $\mathbf{X}$ is the set of class labels on the underlying image of a p-dimensional random field $\mathbf{Y}$ on $\mathcal{S}$. We suppose that $\mathbf{X}$ is comprised of pixels which belong to one of the m classes. For the purposes of the present paper we define a set number of tissue classes. The statistical parameters representing each class are unknown and need to be established at each iteration. Let c denote a clique and $\mathcal{C}$ be the set of all cliques on $\mathcal{S}$. According to the Hammersley-Clifford Theorem, the joint pdf of X is a Gibbs distribution of the form

$$f(X) = \frac{1}{Z} e^{-U(X)}, \qquad Z = \sum_{X} e^{-U(X)}, \tag{1}$$

Z being the partition function and $U(X)$ an energy function defined by:

$$U(X) = \sum_{c \in \mathcal{C}} V_c(X) = \sum_{s \in \mathcal{S}} \sum_{j \in \mathcal{N}_s} \beta(1 - \delta(x_s - x_j)), \qquad \forall s, s' \in c, s \neq s' \tag{2}$$

with $V_c(X)$ being the potential function whose argument, X, is an element of the clique. $\mathcal{N}_s$ denotes the neighborhood pixels of x_s in the associated clique. The MRF we consider here is a multilevel logistic model that has a second order neighborhood system with pairwise cliques: where x_s is the realization of the 1-D random field $\mathbf{X}$ on $\mathcal{S}$. β can be interpreted as edge penalty. The observed

image $\mathbf{y}_s$ is obtained when the noise $\mathbf{w}_s$ is superimposed on the signal $g(\mathbf{x}_s)$:

$$\mathbf{y}_s = g(\mathbf{x}_s) + \mathbf{w}_s \tag{3}$$

where $g(\mathbf{x}_s)$ is a function that maps the underlying label $\mathbf{x}_s$ to its associated attribute $\underline{\mu}_{\mathbf{x}_s}$ and Σ_{x_s}. The $\mathbf{w}_s$s' are independently distributed Gaussian random vectors with zero mean and unknown covariance matrix, which is class conditional. Therefore, the density of $\mathbf{Y}$, given the underlying true image $\mathbf{X} = X$, is

$$f(Y|X) = \prod_{s\in\mathcal{S}} f(\underline{y}_s|x_s). \tag{4}$$

Based on the observed image Y, the problem is to classify the observed random vector $\underline{y}_s$ into one of the m different classes, subject to iteratively estimated parameters for the multivariable Gaussian distribution of the conditional pdf at every position. Based on the assumption that the pixel intensity y_s follows a Gaussian distribution with parameters $\theta_s = \{\mu_\ell, \sigma_\ell\}$, given the class label $x_s = \ell$,

$$p(y_s|x_s) = g(y_s;\theta_\ell) = \frac{1}{\sqrt{2\pi\sigma_\ell^2}} \exp\Big(-\frac{(y_s-\mu_\ell)^2}{2\sigma_\ell^2}\Big). \tag{5}$$

and the conditional independence assumption of $\mathbf{y}$, equation (4), the joint likelihood probability used for segmentation becomes:

$$f(Y|X) = \frac{1}{(2\pi)^{Np/2}} \cdot e^{-\frac{1}{2}\sum_{s\in\mathcal{S}}\left[(y_s-\mu_{x_s})^T(\Sigma_{x_s})^{-1}(y_s-\mu_{x_s})+\ln(\|\Sigma_{x_s}\|)\right]} \tag{6}$$

where $\underline{y}_s$ and $\underline{\mu}_{x_s}$ are p-dimensional vectors while Σ_{x_s} is a p x p matrix and N is the total number of pixels in the image. $\underline{\mu}_{x_s}$, Σ_{x_s} are the mean and covariance associated with each class. Equation (6) can be written in the form $f(Y|X) = \frac{1}{Z'}e^{-U(Y|X)}$ with likelihood energy

$$U(Y|X) = \sum_{s\in\mathcal{S}} U(y_s|x_s) = \frac{1}{2}\sum_{s\in\mathcal{S}} \left[\mathcal{H}(y_s,x_s) + \ln(\|\Sigma_{x_s}\|)\right], \qquad \mathcal{H}(y_s,x_s) = (y_s-\mu_{x_s})^T(\Sigma_{x_s})^{-1}(y_s-\mu_{x_s}). \tag{7}$$

Since $\ln f(\mathbf{x}|\mathbf{y}) = -U(\mathbf{x}|\mathbf{y}) + C$ for some constant C, according to the MAP criterion the segmentation is given as:

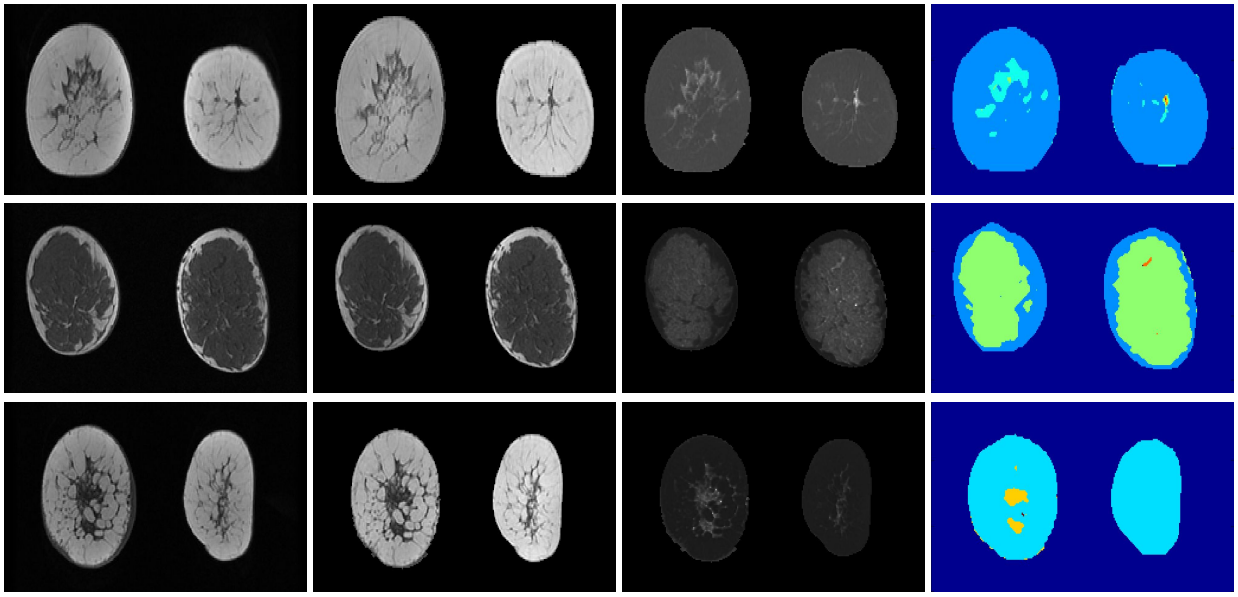
$$\hat{x} = \arg\max_{\mathbf{x}\in\mathbf{X}} f(\mathbf{x}|\mathbf{y}) = \arg\min_{\mathbf{x}\in\mathbf{X}}(U(\mathbf{y}|\mathbf{x}) + U(\mathbf{x})) = \arg\min_{\mathbf{x}\in\mathbf{X}} \sum_{s\in\mathcal{S}} \left[\frac{1}{2}\mathcal{H}(y_s,x_s) + \ln(\|\Sigma_{x_s}\|) + \sum_{j\in\mathcal{N}_i} \beta(1-\delta(x_s-x_j))\right]. \tag{8}$$

Segmentation/classification is the result of minimizing the solution of equation (8) iteratively. Our solution is given via the SA algorithm that requires an initial segmentation of the image and for that purpose we apply the maximum likelihood segmentation as the first step of the process. The SA algorithm overcomes local minima and provides the unique global minimum that corresponds to the correct segmentation. This is achieved by using random sampling by allowing an occasional increase in the value of an the associated function included in brackets of equation (8). In conjunction with stochastic annealing that forces samples of the posterior distribution towards the minimal global energy configuration, local minima are avoided and convergence to the global minimum is guaranteed [7].

Given the data $\mathbf{y}$, the annealing procedure begins with an initial configuration $x_{\mathcal{S}}^{(0)}$ that is chosen to be the maximum likelihood segmentation in this application, the initial configuration is assigned a very high temperature T, so that all configuration changes can be accepted. New configurations are drawn successively from a sampling process with T dropping down gradually until minimum temperature is reached. By choosing carefully the decreasing sequence of temperatures, the convergence to the global minimum is guaranteed [7]. The SA algorithm considers, instead of the joint likelihood pdf, the pdf $\log f_T(\mathbf{x}|\mathbf{y}) \cong -\{U(\mathbf{y}|\mathbf{x})U(\mathbf{x})\}^{\frac{1}{T}}$, and as $T \to 0$ sequentially, given by $T^{(t)} = kT^{(t-1)}$ with $T^{(0)} = 1$ and $k = 0.9$, the algorithm sequentially updates each $x_s^{(k)}$ into $x_s^{(k+1)}$ by minimizing $U(x_s|\mathbf{y}, x_{\mathcal{S}-\{s\}})$, the conditional posterior probability, with respect to x_s. After the initial classification, we evaluate the class means and covariances needed for the SA optimization method using the following equations:

$$\underline{\hat{\mu}_l} = \frac{1}{r_l}\sum_{s\in\mathcal{S}} \underline{y_s}, \qquad \underline{\hat{\Sigma}_l} = \frac{1}{r_l}\sum_{s\in\mathcal{S}} (\underline{y_s}-\underline{\mu_l})(\underline{y_s}-\underline{\mu_l})^T, \quad x_s = l \tag{9}$$

where $l = 1,\ldots,m$ and r_l is the number of points in class l. The method continues by iterating the SA algorithm followed by parameter updating using equations (9) until the segmentation converges.



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

Figure 1. (a) Original breast MR slice. (b) Breast MR slice after automatic removal of the bias field. (c) The T_1 map as an intensity image corresponding to the breast MR slices at (b). (d) The segmentation result using 8 tissue classes after 50 iterations of the SA algorithm.

4 Results, Discussion and Conclusion

We have applied the presented HMRF to the segmentation of T_1 parametric maps resulting from clinical MR studies of 3 patients and we have obtained excellent segmentation results, depicted in Figure 1. As it can be seen, choosing 8 tissue classes results in good correspondence with breast anatomy. In our examples these tissue classes are subsets of the more general categories of fibroglandular tissue and fat. The segmentation results will be even better when the algorithm is applied to T_1 maps corresponding to contrast enhanced MR images of the breast as tissue differentiation will be superior.

This paper presents a new approach in segmentation of T_1 parametric maps of breast MR images using GHMRFs to impose contextual constraints on image pixels. Segmentation is the result of a MAP estimation applied on the T_1 parametric map and obtained using SA. Such a segmentation approach achieves a more physiologically plausible characterization of contextual constraints between neighboring pixels and results in a more reliable real time segmentation than conventional methods. The illustrated results, obtained after applying the method to 3 patients, capture reliably the anatomical information contained in the original images and demonstrate that the method works in practice. The algorithm is fast and the segmentation steps are completely automated. We are now evaluating a version of this algorithm where the number of tissue classes is allowed to change and where the GHMRF parameters are updated using an EM method. Our approach will be used in 2-D registration and reconstruction of the breast. In the future, we aim to create statistical models of T_1 for healthy, malignant and fat tissues.

References

1. G. Ketsetzis & M. Brady. "A new protocol for contrast-enhanced MRI of the breast." In *7th International Conference in Digital Mammography, Durham, USA*, p. 10. Medical Physics Press, 2004.
2. M. Styner, G. Gerig, C. Brechbhler et al. "Parametric estimate of intensity inhomogeneities applied to MRI." *IEEE Transactions on Medical Imaging* **19(3)**, pp. 153–165, 2000.
3. L. P. Clarke, R. P. Velthuizen, M. A. Camacho et al. "MRI segmentation: methods and applications." *Magnetic Resonance Imaging* **13(3)**, pp. 343–368, 1995.
4. S. C. Zhu & A. Yuille. "Region competition: Unifying snakes, region growing, and bayes/MDL for multiband image segmentation." *IEEE Trans. Pattern Anal. Machine Intell.* **18(9)**, pp. 884–900, 1996.
5. Y. Zhang, M. Brady & S. Smith. "Segmentation of brain MR images throught a hidden markov random field model and the expectation maximazation algorithm." *IEEE Trans. Med. Imag.* **20(1)**, pp. 45–57, 2001.
6. J. K. Fwu & P. M. Djuric. "Unsupervised vector image segmentation by a tree structure-ICM algorithm." *IEEE Transactions on Medical Imaging* **15(6)**, pp. 871–880, 1996.
7. S. Geman & D. Geman. "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images." *IEEE Trans. Pattern Anal. Machine Intell.* **6**, pp. 721–741, 1984.

2 Dimensional Electrophoresis Gel Registration Using Point Matching and Local Image-Based Refinement

Mike Rogers[a], Jim Graham[a] and Robert P. Tonge[b]
[a]University of Manchester. [b]Protein Science, Enabling Science and Technology (Biology), AstraZeneca.

Abstract. We present a robust and accurate 2-DE gel alignment algorithm which combines point matching and local image-based refinement. The algorithm uses a novel combination of Euclidian, shape context, image and feature based attributes to produce a point distance measure. Correspondence is determined using this measure and is further improved using an iterative M-estimation approach, and shown to be robust in the presence of large image distortions. Local image-based refinement is shown to improve significantly alignment accuracy. The high accuracy and robustness of the resulting system indicates that it is a promising method for use in practical gel alignment situations.

1 Introduction

Recently, proteomics research has become a large growth area in the bio-sciences. Often, studies involve differential analysis of large sets of 2-D Electrophoresis (2-DE) gels. 2-DE is a method of protein separation that results in a matrix of diffuse spots which can be visualised by pre or post staining. Each of these spots is a separated protein strain. The volume of each spot is proportional to the amount of each protein in the original sample. Figure 1 shows an example of a section from such a gel. To carry out a differential investigation, it is necessary to determine correspondence between spots on sets of gel images. This implies that a transformation relating one gel image to another is required. The production of 2-DE gels is inherently variable. As a result complex non-linear deformations are often required to align comparable gels. These deformations are often difficult to identify manually and time-consuming to correct. The goal of this work is to develop an image registration scheme that can be used to bring pairs of gels into alignment. However in implementing such a scheme, it is important to retain genuine sample differences between the gel pair.

2 Method

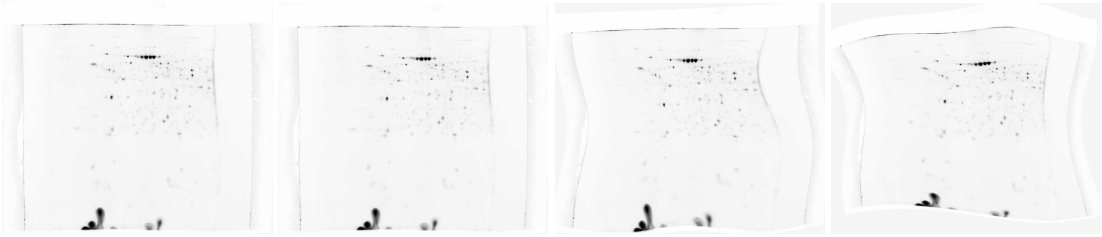
Our gel registration algorithm seeks to combine the global correspondence properties of point matching schemes, with the good local refinement properties of image based registration. First protein spot centre point features are extracted from a pair of images. These points are corresponded using a distance metric combining Euclidean distance, the context of neighbouring spot positions and the context of the distribution of spot intensity and size. The transformation parameters arising from this match are calculated and refined using local image correlation. The resulting transformation is used to initialise and constrain the next point matching step. The process is iterated until a convergence criterion is met. The stages of our algorithm are described in more detail in sections 2.1-2.4. Throughout the process we have used Clamped-Plate Splines (CPSs) [1] to parameterise non-rigid transformations, after extracting any affine component. CPSs are interpolating spines, similar to Thin-Plate Splines, but they use an alternative Green's function that yields improved boundary conditions on the unit circle and can be converted into weighted smoothing splines in the standard manner. The whole algorithm has been implemented in a multi-resolution framework, with the final transform from the current resolution being used to initialise the next highest. To help avoid local minima we have varied the amount of regularisation applied during spline calculation at each resolution. Starting with a strongly constrained smooth transform at coarse resolutions, the smoothing parameter, $0 \geq p \leq 1$ is decreased at each resolution level, ending with a less constrained transform. In this work we vary p linearly between 0.25 and 0.01. We now describe each stage of our algorithm in more detail.

2.1 Feature Extraction

We have used an extremely simple point feature extraction process to detect spot centres at each resolution. We are able to do this as the rest of the algorithm has been designed to be robust to large numbers of outliers and noise in the original point sets. To calculate a binary image feature image, we use a threshold on the Laplacian of each image as follows: $f = (\partial_x^2 > t) \wedge (\partial_y^2 > t)$, where ∂_x^2 and ∂_y^2 are the 2nd derivative of the Gaussian smoothed gel image in the x and y directions respectively and t is a threshold. We have chosen to smooth with Gaussian $\sigma = 2$ pixels and threshold at $t = 0$. The centre of gravity of each connected region in the image meeting this criterion is taken as a point location. We have limited the number of points to a maximum by discarding all but the 400 most intense spots, measured using image intensity information. The number of features to retain was chosen manually. Our spline transformation model therefore has a maximum of 400 control points.

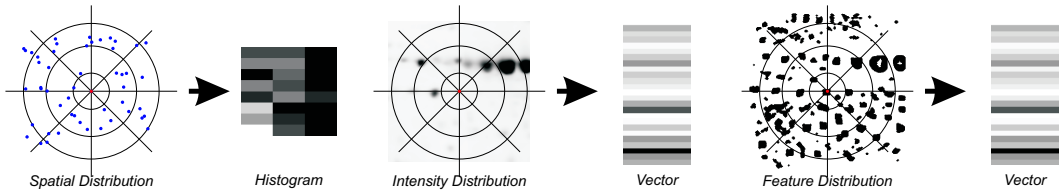
2.2 Point Matching

In previous work, we have developed a point matching strategy suitable for use when aligning gel images [2]. In this work, we have extended the basic scheme in several ways, and we will now briefly describe the algorithm



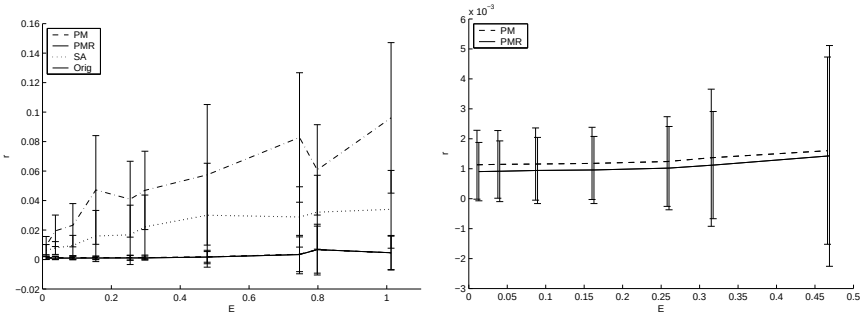
(a) Original (b) Small deformation (σ=0.005) (c) Medium deformation (σ=0.025) (d) Large deformation (σ=0.05)

Figure 1. Examples of 2-DE gel image. (b)-(d) have been synthetically deformed.



(a) Shape Context (b) Image Context (c) Feature Context

Figure 2. Attributes for feature distance calculation.



(a) (b)

Figure 3. (a) Mean residual r for each algorithm (PM,PMR,SA) and original deformation (Orig) with ± 1 std.dev. error bars. r plotted as a percentage of image size. x-axis shows mean deformation energy $\bar{E}$. (b) Same data, showing only first 7 measurements for PM and PMR (error bars slightly offset for display).

together with our extensions. Our point matching scheme is based on the commonly used ICP algorithm [3]. In [2] we used the Shape Context (SC) [4] measure as an alternative distance metric to Euclidian distance. SC provides a semi-global description of the spatial distribution of neighbouring points by counting the number of points in radial regions, yielding histograms that can be made invariant to affine deformations (see figure 2(a)). The method also includes an explicit treatment of outliers. The χ^2 statistic between histograms is used as a distance between features. Careful evaluation in the presence of outlier features has shown that when deformation is expected to be large the most appropriate distance measure is SC, and when deformation is small Euclidian (Euc) distance yields the highest accuracy and robustness. In addition to the distance measure presented in [2], we have added two more attributes. Following the SC histogram binning method, we have developed semi-global image intensity and feature information descriptors. As illustrated in figures 2(b) and 2(c), rather than counting the number of points in a specific bin, we use the average image intensity within the region to form one element of an attribute vector. The Euclidian distance between these vectors could be calculated as a feature distance measure. However, we know there will be important genuine differences between the intensities and patterns of spots in two gels. This makes Euclidian intensity distance an inappropriate measure. Instead, we have the robust Least Median of Squares (LMedS) measure to calculate the distance between vectors. Using this scheme, we produce two additional distance measures, one associated with the original image intensities (Image Context (IC), figure 2(b)) and another associated with the binary feature image (Feature Context (FC), figure 2(c)) calculated during feature extraction. These measures are combined into a single distance between features using the following formula (neglecting

normalisation): $d' = \alpha d_{Euc} + (1-\alpha)(d_{SC} + d_{IC} + d_{FC})/3$, where d_{IC} and d_{FC} are distances calculated using LMedS, α is a weighting factor between the two measures. All measures are normalised over the set of all distances to have mean 0 and standard deviation 1, which ensures equal influence for each measure. Using a closest point method with this distance measure, correspondence can be determined between feature points. Corresponding points are used as control points to estimate CPS parameters. Due to genuine differences in spot pattern and our basic automatic feature extraction scheme we know that correspondences will contain errors. We have used the iterative M-estimation paradigm to down-weight correspondences that are inconsistent with their neighbours. M-estimation calculates weights for each data point based on their residual distance against a model. In our case, residuals are calculated as the Euclidian distance between the predicted position of the feature given by the regularised CPS transform and the associated corresponding point position. We have used the Huber kernel [5] to weight the correspondences. CPS parameters are re-calculated from the weighted correspondences. The process is iterated until convergence.

2.3 Local Image-Based Refinement

A further refinement to [2] addresses the inaccuracy of feature localisation using the centre of gravity of regions within the binary feature image. The centre of gravity of corresponding feature areas in two different gels may not be in the same position on the gels. For this reason, we optimise the position of each point in one image w.r.t. the location of corresponding point in the other. A simple gradient descent process minimises the cross-correlation between local image patches centred at the location of each corresponding pair of points, by adjusting the location of a point in one of the images. The new feature locations are used in subsequent feature matching iterations.

2.4 Convergence Criterion

The criterion chosen to determine algorithm convergence is difficult to define. At present we simply use the mean weighted Euclidian distance between corresponding features (determined in 2.2). If the difference between this value at successive iterations is less than 10^{-6} then the algorithm is said to have converged.

3 Evaluation

We have evaluated our point matching algorithm, both with and without local image refinement, in comparison with the softassign approach described in [6]. As subsequent analysis of 2-DE gels requires the comparison of corresponding protein spots, the effectiveness of gel matching algorithms should be measured in terms of the accuracy of alignment of protein spots. To perform this evaluation we require a large set of gel image pairs with annotated spot positions and known correspondence. Ideally, matching difficulty for each pair should be known and should represent the true range found in real data. Data meeting these requirements is not available and, due to the complexity of the images, would be extremely time consuming to produce. Instead, we have used DIGE gel pairs with known spot locations and introduced varying amounts of synthetic deformation to form our test data set. DIGE gels [7] are produced using protein mixtures that are pre-stained with different fluorescent dyes. The dyes are chosen to fluoresce under different frequencies of UV light. After staining, up to three samples can be mixed together and run on a single gel. Corresponding proteins from different samples will migrate to exactly the same place on the gel and be coincident in gel images. To retrieve images from the separate samples, the gel is illuminated with the excitation frequency for each of the dyes, allowing pairs of images to be produced with perfect correspondence but showing genuine sample differences. In this evaluation, we have used 5 pairs of DIGE gel images, each with $\sim$650 annotated spot positions. Using these images, we created a large evaluation data set by introducing varying amounts of synthetic deformation to one of the images of each pair. The amount of deformation has been controlled as follows: 10 control points are sampled from a uniform random distribution. 10 random offsets are sampled from a Gaussian with known σ. A smooth Gaussian RBF transformation ($p = 0.05$) is calculated using the control points and offsets and used to transform a gel image and its spot positions. Increasing the value of σ increases the amount of deformation. An estimate of the deformation energy E can be calculated from the parameters of the RBF ($E = \sum \mathtt{diag}(\vec{a}' A \vec{a})$). Figure 1 shows examples of images deformed using different values of σ. In our evaluation, σ has been varied linearly in 10 steps between $0.005 \rightarrow 0.05$ (E : $0.0061 \rightarrow 1.2$). By observation, the top end of this range greatly exceeds the maximum amount of deformation required to align corresponding gels in practice. At each value of σ, we have created 5 randomly deformed images from each DIGE pair. This gives a total of $5 \times 5 \times 10 = 250$ gel alignments, each with $\sim$650 spots. After gel alignment, the recovered transformation is used to transform the spot locations to their estimated position in the un-deformed gel. We have measured the residual Euclidian distance between the transformed spots and their ground-truth position (r). Residual r is reported as a proportion of the maximum dimension of the associated gel image. In this way, $r = 0.01$ represents an error of 1% of image size.

4 Results and Discussion

Figure 3(a) shows mean and standard deviation of r after alignment for each algorithm: Point Matching with Refinement (PMR), Point Matching (PM) and SoftAssign (SA). The results are plotted against mean deformation

$\bar{E} \approx$	0.01	0.08	0.25	1.02
PM	0.15	0.015	0.26	11.29
PMR	0.02	0.013	0.21	10.28
SA	64.93	71.33	84.1	93.77

(a) $r > 1\%$ image size

$\bar{E} \approx$	0.01	0.08	0.25	1.02
PM	0	0.01	0.04	8.19
PMR	0	0.01	0.07	7.72
SA	15.62	30.94	48.54	79.35

(b) $r > 2\%$ image size

Table 1. Percentage of residuals greater than 1% and 2% of image size.

energy ($\bar{E}$) over replicates at each value of σ. A residual value of $r = 0$ indicates a perfect recovery of the alignment of a protein spot. Results for the set of 5 images have been combined and as such each data point represents $\sim$16250 point residuals. Also shown in figure 3(a) are values for un-aligned point residuals (Orig) showing the amount of deformation in the original data in terms of point residual. The data for PM and PMR are almost coincident on the scale of figure 3(a). Figure 3(b) shows the same data, this time including only the first 7 groups for PM and PMR. Large error bars on groups $8 \rightarrow 10$ prevent easy visual comparison of group means, and so have not been displayed on this figure. Each algorithm produces reduced residuals compared to the original deformations. However, Both PM and PMR are significantly more accurate than SA. SA is a state-of-the-art general purpose global point matching scheme. In this case, our carefully designed task-specific multi-resolution local refinement scheme performs significantly better. The SA scheme in [6] would have to be significantly redesigned in a multi-resolution framework to obtain similar accuracy as our PM or PMR algorithms. Figure 3(b) shows that PMR, which uses local image refinement, produces consistently lower residuals than PM. The reduction in residual mean is significant at the 0.01 level for all groups, except when $\sigma = 0.045 (\bar{E} \approx 0.75)$. The mean residual achieved by both methods is almost constant up to a bending energy of $\bar{E} \approx 0.25$, showing robustness to deformations up to this magnitude. We have observed that this is a realistic upper level for the amount of deformation required to align most 2-DE gel pairs, however further work is required to validate this. PM and PMR give a residual value of $r \approx 10^{-3}$ for the $\bar{E} \approx 0.25$ group, which corresponds to a mean protein spot location residual of 1 pixel in a 1000×1000 pixel image. For a gel registration system to be of use in practice, alignment must be accurate across the entire gel. If groups of misaligned spots are present manual validation and correction of results would be required. This is a time-consuming and subjective process which should be avoided. We have evaluated the numbers of large point residuals by counting the number of residuals greater than a threshold. Tables 1(a) and (b) show the percentage of residuals larger than 1% and 2% of image size for a selection of values of $\bar{E}$. For users to have high confidence in a gel alignment system, very few large residuals must be produced. Using 1% of image size (10 pixels in a 1000×1000 pixel image) as a threshold, both PM and PMR produce fewer than 45 large residuals out of ~ 16500 measurements. Using the 2% threshold both produce less than 10 large residuals. In contrast, SA results in ~ 780 and ~ 1400 respectively. The small numbers of large residuals produced demonstrates that either PM or PMR may be suitable for use for automatic gel alignment.

5 Summary

We have presented a robust and accurate point matching based 2-DE gel alignment algorithm. The algorithm uses a novel combination of Euclidian, shape context, image and feature based attributes to produce a point distance measure capable of determining good correspondence between protein spot point sets. This correspondence is further improved using an iterative M-estimation approach. Adding a local refinement step based on image intensities has been shown to improve significantly alignment accuracy. Our algorithms have been shown to out-perform the softassign approach described in [6]. The high accuracy of the system together with the small number of large spot alignment errors indicates that this system shows promise for use in practical gel alignment situations.

References

1. C. Twining, S. Marsland & C. J. Taylor. "Measuring geodesic distances on the space of bounded diffeomorphisms." In *BMVC*, volume 2, pp. 847–856. 2002.
2. M. Rogers, J. Graham & R. P. Tonge. "2-D electrophoresis gel registration using feature matching." In *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 1436–1439. Arlington, USA, April 2004.
3. P. J. Besl & H. D. McKay. "A method for registration of 3-D shapes." *IEEE Transactions on Pattern Analysis and Machine Vision* **14(2)**, pp. 239–256, 1992.
4. S. Belongie, J. Malik & J. Puzicha. "Shape matching and object recognition using shape context." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24(4)**, pp. 509–522, April 2002.
5. P. J. Huber. *Robust Statistics*. Wiley, New York, 1981.
6. H. Chui, A. Rangarajan, J. Zhang et al. "Unsupervised learning of an atlas from unlabeled point-sets." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26(2)**, pp. 160–173, February 2004.
7. R. Tonge, J. Shaw, B. Middleton et al. "Validation and development of flourescence two-dimensional differential gel electrophoresis proteomics technology." *Proteomics* **1**, pp. 377–396, 2001.

Whole Brain Voxel-based Analysis Using Registration and Multivariate Statistics

C.E. Thomaz[1], J.P. Boardman[2], D.L.G. Hill[3], J.V. Hajnal[2],
A.D. Edwards[4], M.A. Rutherford[2], D.F. Gillies[1], and D. Rueckert[1]

[1]Department of Computing, Imperial College, London, UK
[2]Imaging Sciences Department, Imperial College, London, UK
[3]Division of Imaging Sciences, King's College, London, UK
[4]Division of Paediatrics, Obstetrics and Gynaecology, Imperial College, London, UK

1 Introduction

Whole brain voxel-based morphometry and statistical pattern recognition methods have been used to classify and describe anatomical structures of MR images. Most of these methods are based on statistical learning techniques applied to either segmented images or a number of features pre-selected from specific image decomposition approaches. Although such pre-processing strategies have overcome the difficulty of dealing with the inherent high dimensionality of 3D brain image data, most of these approaches rely on optimisation techniques that are time consuming and do not provide a simple way of mapping the classification results back into the original image domain for further interpretation.

In this paper, we use the general multivariate statistical methodology (PCA+LDA) to identify the most discriminating hyper-plane separating two populations. We introduce some novel techniques to overcome the well-known instability of the LDA within-class scatter matrix and increase the computational efficiency of the approach. Our goal is to analyse all the data simultaneously rather than feature by feature. The result is an efficient and practical method for separating two populations and visually analysing their differences.

2 Methodology

Before we can analyse the MR images we need to map all images into a common atlas coordinate system. This pre-processing step is essential because the construction of the multivariate statistical model relies on anatomical correspondences when comparing patterns across subjects. We have randomly chosen the image of one subject as reference or atlas. In order to map the anatomy of each subject into the anatomy of the atlas we have first applied an affine registration [1] followed by non-rigid registration based on free-form deformations [2]. Both algorithms are based on the maximisation of normalised mutual information as a voxel-based similarity measure.

2.1 PCA

After registration, the Principal Components Analysis (PCA) technique is performed. PCA is a feature extraction procedure concerned with explaining the covariance structure of a set of variables through a small number of linear combinations of these variables. It is a common statistical technique that has been used in several image recognition problems, especially for dimensionality reduction.

Although there is always the question of how many principal components to retain in order to reduce the dimensionality of the original training sample, Yang and Yang [3] have proved recently that the number of principal components to retain for a best LDA classification performance should be equal to the rank m of the total covariance matrix S composed of all the training patterns and given by

$$S = \frac{1}{(N-1)} \sum_{j=1}^{N} (x_{i,j} - \bar{x})(x_{i,j} - \bar{x})^T , \quad \textbf{(1)}$$

where $x_{i,j}$ is the *n*-dimensional pattern j from class π_i, N is the total number of samples, and $\bar{x}$ is the grand mean vector given by

$$\bar{x} = \frac{1}{N}\sum_{j=1}^{N} x_{i,j} \ . \qquad \textbf{(2)}$$

The m principal components can then replace the initial n features and the original data set, consisting of N measurements on n variables, is reduced to a data set consisting of N measurements on m principal components. For this representation to make sense in statistical classification problems we are making the assumption that the distributions of each class or group are separated by their corresponding mean differences.

2.2 LDA

The primary purpose of Linear Discriminant Analysis (LDA) is to separate samples of distinct groups by maximising their between-class separability while minimising their within-class variability. Let the between-class scatter matrix S_b be defined as

$$S_b = \sum_{i=1}^{g} N_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \ , \qquad \textbf{(3)}$$

and the within-class scatter matrix S_w be defined as

$$S_w = \sum_{i=1}^{g} (N_i - 1) S_i = \sum_{i=1}^{g} \sum_{j=1}^{N_i} (x_{i,j} - \bar{x}_i)(x_{i,j} - \bar{x}_i)^T \ , \qquad \textbf{(4)}$$

where $x_{i,j}$ is the *m*-dimensional pattern j from class π_i, N_i is the number of training patterns from class π_i, g is the total number of classes or groups, and $\bar{x}$ is the grand mean vector defined in equation (2). The vector $\bar{x}_i$ and matrix S_i are respectively the unbiased sample mean and sample covariance matrix of class π_i.

The main objective of LDA is to find a projection matrix P_{lda} that maximises the ratio of the determinant of the between-class scatter matrix to the determinant of the within-class scatter matrix (Fisher's criterion), that is

$$P_{lda} = \arg\max_{P} \frac{\left| P^T S_b P \right|}{\left| P^T S_w P \right|} \ . \qquad \textbf{(5)}$$

It is a proven result [4] that if S_w is a non-singular matrix then the Fisher's criterion is maximised when the projection matrix P_{lda} is composed of the eigenvectors of $S_w^{-1} S_b$ with at most $(g-1)$ nonzero corresponding eigenvalues. This is the standard LDA procedure.

However, the performance of the standard LDA can be seriously degraded if there are only a limited number of total training observations N compared to the dimension of the feature space m. Since the within-class scatter matrix S_w is a function of $(N-g)$ or less linearly independent vectors, its rank is $(N-g)$ or less. Therefore in the problem under investigation where the number of training patterns is comparable to the number of features, S_w might be singular or mathematically unstable and the standard LDA cannot be used to perform the task of the classification stage.

2.3 MLDA

In order to avoid both the singularity and instability critical issues of the within-class scatter matrix S_w when LDA is used in such limited sample and high dimensional problem, we have proposed a maximum uncertainty LDA-based approach (MLDA) to overcome the instability of the S_w matrix [5]. It is based on the maximum entropy covariance selection method developed to improve quadratic classification performance on limited sample size problems [6].

The proposed method considers the issue of stabilising the S_w estimate with a multiple of the identity matrix by selecting the largest dispersions regarding the S_w average eigenvalue. The following selection algorithm expands only the smaller and consequently less reliable eigenvalues of within-class scatter matrix S_w:

i. Find the Φ eigenvectors and Λ eigenvalues of S_p, where $S_p = S_w / [N-g]$;

ii. Calculate the S_p average eigenvalue $\bar{\lambda}$ using

$$\bar{\lambda} = \frac{1}{m}\sum_{j=1}^{m}\lambda_j = \frac{tr(S_p)}{m};$$

iii. Form a new matrix of eigenvalues based on the following largest dispersion values

$$\Lambda^* = diag[\max(\lambda_1, \bar{\lambda}), \max(\lambda_2, \bar{\lambda}), ..., \max(\lambda_m, \bar{\lambda})];$$

iv. Form the modified within-class scatter matrix

$$S_w^* = S_p^*(N-g) = (\Phi\Lambda^*\Phi^T)(N-g).$$

The maximum uncertainty LDA is constructed by replacing S_w with S_w^* in the standard Fisher's criterion formula described in equation (5). It is a straightforward method that overcomes both the singularity and instability of the within-class scatter matrix S_w when LDA is used in limited sample and high dimensional problems.

3 Experiments

To demonstrate the effectiveness of the approach, we have used a neonatal MR brain data set that contains 67 preterm infants at term equivalent age (mean 29.7, range 24-34 weeks post-menstrual age), and 12 term born controls (mean 39.3, range 36-42 weeks post-menstrual age). Ethical permission for this study was granted by the Hammersmith Hospital Research Ethics Committee and informed parental consent was obtained for each infant. Infants were sedated for the examination but did not require mechanical ventilation at the time of MR imaging. Pulse oximetry, electrocardiographic and televisual monitoring were used throughout the examination which was attended by a paediatrician. A 1.5 T Eclipse MR System (Philips Medical Systems, Cleveland, Ohio) was used to acquire high resolution T1 weighted images (TR=30ms, TE=4.5ms, flip angle = 30o). In addition to conventional T1 and T2 weighted image acquisition, volume datasets were acquired in contiguous sagittal slices (in-plane matrix size 256 x 256, FOV = 25cm) with a voxel size of 1.0 x 1.0 x 1.6 mm3.

We have performed two main tasks: classification and visual analysis. First a training matrix composed of N zero mean n-dimensional image vectors is used as input to compute the PCA transformation matrix. The columns of this n x m transformation matrix are eigenvectors, in eigenvalues descending order. The N zero mean image vectors are projected on the principal components and reduced to m-dimensional vectors representing the most expressive features of each one of the pre-processed n-dimensional image vector. Afterwards, this N x m data matrix is used as input to calculate the MLDA discriminant eigenvector. The most discriminant feature of each one of the m-dimensional vectors is obtained by multiplying the N x m most expressive features matrix by the MLDA linear discriminant eigenvector. An analogous procedure, but in reverse order, has been used to convert any point on the most discriminant space back to its corresponding n-dimensional image vector. More specifically, first we multiply that particular point by the transpose of the linear discriminant vector previously computed, then we multiply its m most expressive features by the transpose of the principal components matrix, and finally we add the average image calculated in the training stage to the n-dimensional image vector.

4 Results

Figure 1 presents the leave-one-out recognition rate (rr) of the two-stage linear classifier using the affine and non-rigid registration algorithms as pre-processing techniques. As expected, the classification results obtained by the non-rigid registration algorithms are higher than the one obtained by an affine transformation, achieving a maximum recognition rate of 97.47% with a control point spacing of 5mm.

Figure 2 highlights the statistical differences between the preterm infants (shown on the top) at term equivalent age and the control group (bottom) mapped back (without the mean) into the image domain. We can see clearly differences in the ventricular system, the posterior limb of the internal capsule, the corpus callosum area, and the inter-hemispheric fissure.

Registration	rr (%)
Affine	91.14
Non-rigid (10mm)	96.20
Non-rigid (5mm)	97.47
Non-rigid (2.5mm)	93.67

Figure 1. Classification results.

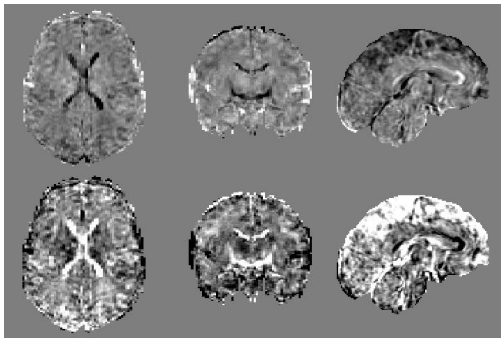


Figure 2. Visual statistical differences.

5 Conclusion

This paper describes the idea of using PCA plus the maximum uncertainty LDA-based approach to classify and analyse MR brain images. The methodology proposed has been performed directly on the MR intensity images rather than on segmented versions of the images. Our results indicate that the use of non-rigid registration in the pre-processing step and the two-stage linear classifier make clear the statistical differences between the control and preterm neonatal samples, showing a classification accuracy of 97.47% using the leave-one-out method.

Although the experiments carried out were based on a specific preterm infants database, we believe that such multivariate statistical strategy for targeting limited sample and high dimensional problems provides a suitable framework for characterising and analysing the high complexity of MR images in general.

Acknowledgements

This work is part of the UK EPSRC e-science project “Information eXtraction from Images” (IXI). Also, the first author was partially supported by the Brazilian Government Agency CAPES under Grant 1168/99-1.

References

1. C. Studholme, D. L. G. Hill and D. J. Hawkes. “An overlap invariant entropy measure of 3D medical image alignment”, *Pattern Recognition*, vol. 32, no. 1, pp 71-86, 1999.
2. D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. “Non-rigid registration using free-form deformations:Application to breast MR images”, *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712-721, 1999.
3. J. Yang and J. Yang, “Why can LDA be performed in PCA transformed space?”, *Pattern Recognition*, vol. 36, pp. 563-566, 2003.
4. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, second edition. Boston: Academic Press, 1990.
5. C. E. Thomaz. “Maximum Entropy Covariance Estimate for Statistical Pattern Recognition”, *PhD Thesis*, Department of Computing, Imperial College London, 2004.
6. C. E. Thomaz, D. F. Gillies and R. Q. Feitosa. "A New Covariance Estimate for Bayesian Classifiers in Biometric Recognition", *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics*, vol. 14, no. 2, pp. 214-223, February 2004.

Computing Covariances for "Mutual Information" Coregistration

P. A. Bromiley*, M. Pokric, and N.A. Thacker

Imaging Science and Biomedical Engineering, Stopford Building,
University of Manchester, Oxford Road, Manchester, M13 9PT.

Abstract. Mutual information (MI) has become a popular similarity measure in multi-modality medical image registration since it was first applied to the problem in 1995. This paper describes a method for calculating the covariance matrix for MI coregistration. We derive an expression for the covariance matrix by identifying MI as a biased log-likelihood measure. The validity of this result is then demonstrated through comparison with the results of Monte-Carlo simulations of the coregistration of T1-weighted to T2-weighted synthetic MRI scans of the brain. We conclude with some observations on the theoretical basis of MI as a log-likelihood.

1 Introduction

The use of MI as a similarity measure for multi-modality coregistration was first proposed in 1995 [1], and it has since become the most popular information-theoretic approach to this problem. Research into coregistration has generally focused on the definition of similarity metrics or on the representation of the transformation model. There is however a growing recognition that characterisation of the accuracy of coregistration is essential if further quantitative processing of the images is to be performed using the resultant transformation model. For example, Crum et. al. [2] state that "...the veracity of studies that rely on non-rigid registration should be keenly questioned when the error distribution is unknown and the results are unsupported by other contextual information". We present an analytical expression for the covariance matrix of the parameters of MI coregistration, based on the identification of the measure as a biased log-likelihood. This is only the first step towards a full characterisation of the error for the general coregistration problem: for example, it takes no account of the difference between image similarity and biological correspondence. It does however provide a lower bound on the error, which may be achievable for certain coregistration problems and definitions of correspondence.

Mutual information $\mathcal{I}(I;J)$ measures the Kullback-Leibler divergence [3] between the joint probability distribution $p(i,j)$ of two images or image volumes I and J and the product of their marginal distributions $p(i).p(j)$ [3],

$$\mathcal{I}(I;J) = \sum_{i,j} p(i,j) \log \frac{p(i,j)}{p(i).p(j)}$$

i.e. the divergence of the joint distribution from the case of complete independence of the images, where the sum is performed over a joint intensity histogram of the image pair. Therefore, maximisation of this measure with respect to a set of coregistration parameters will optimise the image alignment. Following [4], we can split the sum into

$$\mathcal{I}(I;J) = \sum_{i} p(i) \log \frac{1}{p(i)} + \sum_{i,j} p(i,j) \log \frac{p(i,j)}{p(j)}.$$

Recognising that the first term on the R.H.S. is the entropy $H(I)$ of image I [3] and that in the limit of large samples $p(i,j) = N_{ij}/N$, where N_{ij} is the number of entries in histogram bin (i,j) and N is the total number of entries in the histogram, we obtain

$$N[\mathcal{I}(I;J) - H(I)] = \sum_{v} \log \frac{p(i,j)}{p(j)} = \log P(I|J),$$

where v represents a sum over voxels rather than histogram bins, and P represents a probability summed over an entire pair of images or volumes. At this point we can make the arbitrary definition that I is the target (fixed) image and J the source image i.e. the image altered by the transformation model. If we ensure that the data sampled from the target image does not change by keeping the overlap with the source image constant[1], $H(I)$ will be a constant, indicating that MI is then a monotonic function of the log-probability of image I given image J,

$$logP(I|J) = N(\mathcal{I}(I;J)) + const. \tag{1}$$

*E-mail: paul.bromiley@talk21.com

[1]Excluding an appropriately sized border around the target image will ensure that all of the remaining data overlaps the source image throughout the optimisation.

The covariances for a maximum-likelihood (ML) method are bounded by the minimum variance bound (MVB) [5]

$$C^{-1}_{\theta_{r,s}} = -\frac{\partial^2 \log L}{\partial\theta_r \partial\theta_s}\bigg|_{\theta_0}$$

where the θ represent parameters of some model, L represents the likelihood function, and θ_0 represents the parameters at the maximum of L. This bound becomes exact if the log-likelihood is quadratic i.e. the likelihood function is Gaussian. Proceeding with the Gaussian assumption, we can write

$$L = \prod_d A_d e^{-\frac{(I_d - I_M)^2}{2\sigma_d^2}} \quad \Rightarrow \quad \log L = \sum_d -\frac{(I_d - I_M)^2}{2\sigma_d^2} + logA_d \quad \Rightarrow \quad \frac{\partial^2 \log L}{\partial\theta_r \partial\theta_s}\bigg|_{\theta_0} = \sum_d -\frac{1}{\sigma_d^2}\frac{\partial I_M}{\partial\theta_r}\frac{\partial I_M}{\partial\theta_s}\bigg|_{\theta_0}$$

where A_d is the normalisation of the Gaussian, I_d are the data and I_M the corresponding model predictions, and σ_d is the standard deviation of the data. Note that any constant normalisation of the Gaussian (A_d) disappears upon differentiation. In simple ML techniques e.g. linear least-squares fitting, the normalisation of L will indeed be constant. However, the MI measure is a "bootstrapped" likelihood, constructed from the joint histogram rather than from some explicit model and so the usual normalisation (to the area under the distribution) may no longer be constant: for example, simply altering the histogram bin size will alter the normalisation. Fortunately, a solution is available in the form of the χ^2 metric. If we normalise to the peak of the distribution, A_d becomes 1 and disappears upon taking logs. Maximisation of the log-likelihood is then directly equivalent[2] to minimisation of the χ^2

$$\log L = \sum_d -\frac{(I_d - I_M)^2}{2\sigma_d^2} = -\frac{\chi^2}{2}.$$

Whilst this is explicitly true for a Gaussian L, we would suggest that this statistic has higher utility regardless of the form of the underlying distribution as it provides appropriate normalisation. Furthermore, the χ^2 can be written in terms of a sum over individual data terms: the so-called χ of the χ^2

$$\chi^2 = \sum_d \chi_i^2 = \sum_i -2\log(L_d) \Rightarrow \chi_d = \sqrt{-2\log L_d}.$$

The expression for the MVB can also be rewritten in this form, through comparison with the previous result for a Gaussian likelihood function

$$\chi_d = \frac{(I_d - I_M)}{\sigma_d} \quad \Rightarrow \quad \sum_d \frac{\partial\chi_d}{\partial\theta_r}\frac{\partial\chi_d}{\partial\theta_s} = \sum_d \frac{1}{\sigma_d^2}\frac{\partial I_M}{\partial\theta_r}\frac{\partial I_M}{\partial\theta_s} \quad \Rightarrow \quad C_\theta^{-1} = \sum_d (\nabla_\theta \chi_d)^T \otimes (\nabla_\theta \chi_d)\bigg|_{\theta_0}. \qquad (2)$$

The Gaussian assumption need only hold over a sufficient range around the maximum to allow the calculation of the derivatives, and since in coregistration the likelihood distribution is composed of tens of thousands of individual data terms (voxels) we can expect, due to the Central Limit Theorem, that this assumption will hold.

The equivalent χ^2 term in the MI measure can be identified using the log-likelihood from Eq. 1

$$logP(I|J) = \sum_v \log\frac{p(i,j)}{p(j)} = \sum_v \log\frac{p(i,j)p(i_{max},j)}{p(j)p(i_{max},j)} = \sum_v \log\frac{p(i,j)}{p(i_{max},j)} + \sum_v \log\frac{p(i_{max},j)}{p(j)}.$$

The first term on the RHS is the χ^2 metric, normalised to the distribution peak $p(i_{max},j)$ as required. The second is a bias term dependent on the non-uniform normalisation of the likelihood distribution. This expression elucidates the behaviour of the MI measure: it is a ML measure biased with a term that maximises the "peakiness" of the distributions in the joint histogram, in order to maximise the correlation between equivalent structures in the images. If we assume that the bias term varies slowly compared to the χ^2 term, which is reasonable since it depends on the marginal distribution, then Eq. 2 can be used. Applying the chain rule to expand the derivative of χ_v (the χ for each data term i.e. voxel) w.r.t. the model parameters in terms of the likelihood for each voxel L_v and the voxel values themselves, J_v, gives

$$C_\theta^{-1} = \sum_v (\frac{\partial\chi_v}{\partial L_v})^2 (\frac{\partial L_v}{\partial J_v})^2 (\nabla_\theta J_v)^T \otimes (\nabla_\theta J_v)\bigg|_{\theta_0}$$

and so

$$C_\theta^{-1} = -\sum_v \frac{(\frac{\partial p(i,j)}{\partial J_v} - \frac{p(i,j)}{p(i_{max},j)}\frac{\partial p(i_{max},j)}{\partial J_v})^2}{2p(i,j)^2 \log\frac{p(i,j)}{p(i_{max}|j)}} (\nabla_\theta J_v)^T \otimes (\nabla_\theta J_v)\bigg|_{\theta_0}. \qquad (3)$$

[2]The χ^2 used here refers to the general definition, as used in statistical tests for assessing the adequacy of fitting results for a number of degrees of freedom, not the specific computational forms used for comparing histograms or tables.

2 Method

The covariance estimation technique was tested on the rigid coregistration of T2 to T1 weighted simulated MR image volumes of a normal brain, obtained from Brainweb [6]. Each volume consisted of 55 slices of 217 by 195 voxels, with Gaussian noise added at 1% of the dynamic range. MI coregistration was implemented within the TINA software package (www.tina-vision.net), using simplex minimisation, allowing the coregistration to optimise the rotation (as Euler angles), translation and scaling of the images. The source images were rotated by 5^o prior to coregistration in order to suppress interpolation artefacts[3], following the suggestion by [7]. However, the coregistrations were started from the correct alignment.

Monte-Carlo simulations were run by adding Gaussian noise to the source image at levels of 0.25 to 2.5 times the original image noise, in ten steps of 0.25σ: covariances were calculated from 1000 coregistrations performed at each noise level. Then Eq. 3 was applied at each noise level, taking the median of 100 estimates of the covariances, over a range around the maximum that represented a change of around 0.5% in the χ^2, in order to stabilise the calculation against the effects of interpolation artefacts, local minima etc. Finally, the estimated and practical (Monte-Carlo) covariances were compared. The covariance matrices were prepared from a set of $1 \times n$ vectors of parameters, and so had only n degrees of freedom despite containing n^2 parameters. It is therefore sufficient to compare only the n diagonal parameters, the variances, or their square-roots, the standard deviations.

3 Results

Fig. 1. shows the dependence of the standard deviations of the transformation model parameters on added noise. The Monte-Carlo results scale linearly with the addition of noise as expected, and linear least-squares fits to the data are shown. Some outliers are present at high noise levels due to bimodality in the Monte-Carlo results: the added noise sufficiently destabilised the coregistration that a local maximum close to the global maximum began to contribute. Therefore, these points were omitted from the fitting process. The estimates from the analytical expression are also shown together with linear fits. The estimated covariances of the translation parameters are identical to the Monte-Carlo results to within the noise on the data. The estimates for the rotational parameters show some divergence, and are also notably noisier, due to the non-linear nature of rotational transformations. The estimates for the scaling parameters show the greatest divergence at the higher noise levels. This is due to an effective underestimate of the covariance through the Monte-Carlo experiments. The scaling parameters are more susceptible to interpolation artefacts than the other parameters, leading to oscillations in the similarity metric around the global maximum. The Monte-Carlo results tend to fall into the local optima generated by these oscillations, leading to underestimation of the covariances, whereas the estimated covariance was stabilised against this effect by taking the median value over 100 points around the global maximum. The sources of the differences between the Monte-Carlo and estimated covariances were identified by plotting the likelihood function around its maximum. Overall, all of the estimated covariances either match the Monte-Carlo results closely, or converge at low noise levels, and are always within a factor of two.

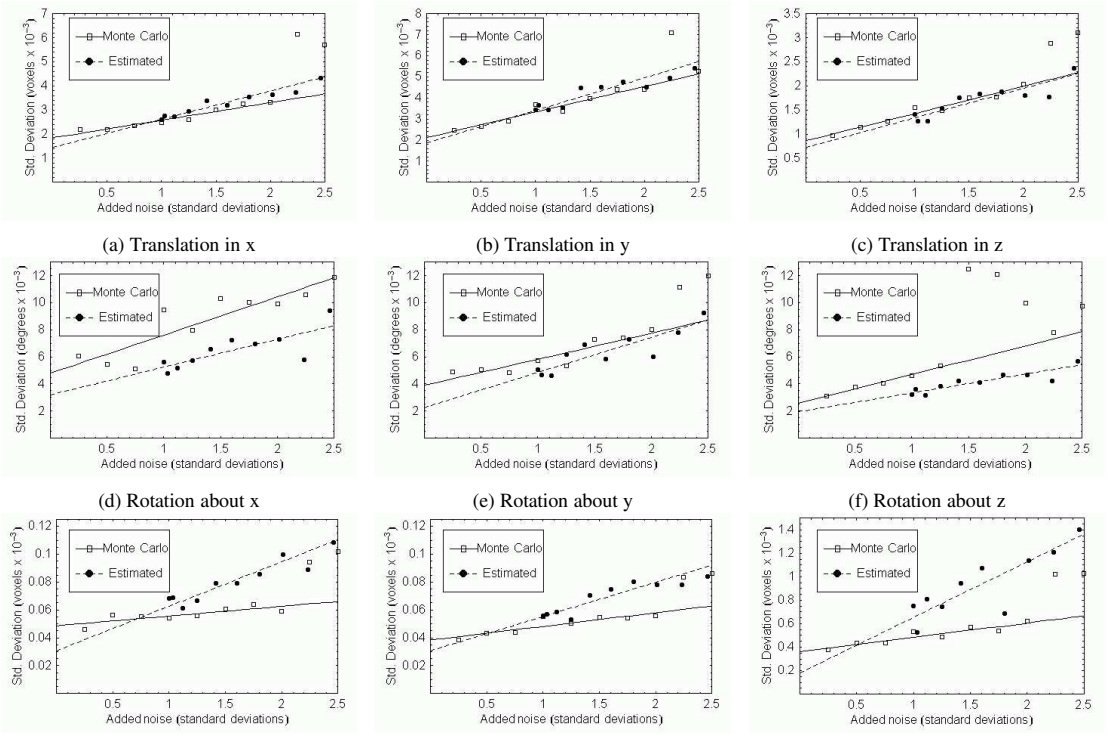
4 Conclusion

This paper has provided a derivation of an analytical expression for the covariances of the parameters of MI coregistration. The result had been confirmed through comparison with Monte-Carlo simulations. The estimated variances match the Monte-Carlo results, both confirming the validity of the covariance estimation technique and justifying the assumption that the MI bias term is negligible in this case.

The covariance estimate presented here has a number of practical uses. Equivalence between estimated and Monte-Carlo variances can be used to demonstrate numerical stability in the implementation of a coregistration algorithm. Error propagation can be used to calculate spatial errors on voxel locations from the covariance matrix of the transformation model. Finally, the technique is equally valid for non-rigid coregistration, where the coregistration errors will vary spatially and so must be quantified if any further statistical analysis of the data is to be performed.

The derivation also illustrates some features of MI in general. Most important is the relationship between MI and log-likelihood. The consistency between the estimated covariances and the practical coregistration performance confirms that this interpretation is valid. We maintain that this is the true theoretical basis of the method, rather

[3] Such artefacts occur due to sudden changes in the number of voxels from the two image volumes that overlap exactly, and so are added to the joint histogram without interpolation. Interpolation smooths the data, so this results in a sudden change in the MI measure.



(a) Translation in x (b) Translation in y (c) Translation in z

(d) Rotation about x (e) Rotation about y (f) Rotation about z

(g) Scaling in x (h) Scaling in y (i) Scaling in z

Figure 1. The standard deviations of the coregistration parameters for the synthetic data. The lines show linear least-squares fits to the data, omitting the outliers from the Monte-Carlo experiments due to evidence of bimodality around the maximum (see main text).

than its relationship to concepts of entropy. The likelihood interpretation may also provide new perspectives on mutual information and associated similarity measures, suggesting alternatives based on quantitative statistics. For instance, normalised MI measures [7] are currently used for coregistration problems with varying sample sizes. The approach adopted here suggests using a χ^2 metric i.e. an appropriately normalised log-likelihood, accommodating the variation in sample size as a change in the number of degrees of freedom. Ultimately, this could lead to a coregistration algorithm implemented in expectation-maximisation form.

Acknowledgements

The authors would like to acknowledge the support of the EPSRC and the MRC (IRC: From Medical Images and Signals to Clinical Information), and of the European Commission (An Integrated Environment for Rehearsal and Planning of Surgical Interventions). All software is freely available from our web site www.tina-vision.net.

References

1. P. Viola & W. M. Wells. "Alignment by maximisation of mutual information." *International Journal of Computer Vision* **24(2)**, pp. 137–154, 1997.
2. W. R. Crum, L. D. Griffin, D. V. G. Hill et al. "Zen and the art of medical image registration: correspondence, homology, and quality." *Neuroimage* **20**, pp. 1425–1437, 2003.
3. T. M. Cover & J. A. Thomas. *Elements of Information Theory*. John Wiley and Sons, New York, 1991.
4. A. Roche, G. Malandain, N. Ayache et al. "Towards a better comprehension of similarity measures used in medical image registration." In *Proceedings MICCAI'99*, pp. 555–566. 1999.
5. R. J. Barlow. *Statistics: A Guide to the use of Statistical Methods in the Physical Sciences*. John Wiley and Sons Ltd., UK, 1989.
6. C. A. Cocosco, V. Kollokian, R. K.-S. Kwan et al. "Brainweb: Online interface to a 3D MRI simulated brain database." *Neuroimage: Proceedings of the 3rd International Conference on Function Mapping of the Human Brain, Copenhagen, May 1997* **5(4)**, pp. S425, 1997.
7. J. P. W. Pluim, J. B. Antoine Maintz & M. A. Viergever. "Interpolation artefacts in mutual information-based image registration." *Computer Vision and Image Understanding* **77**, pp. 211–232, 2000.

Groupwise Non-Rigid Registration of Medical Images: The Minimum Description Length Approach

Carole J. Twining[1]*, Stephen Marsland[2]*, and Chris Taylor[1]
[1] ISBE, University of Manchester, U.K.
[2] Institute of Information Sciences, Massey University, N.Z.

Abstract. The aim of non-rigid registration as applied to a set of images is to find a 'meaningful' dense spatial correspondence across the set. There are many methods available for finding such a correspondence given a *pair* of images, but viewing the groupwise case as repeated successive pairwise is rather naïve – the principled non-rigid registration of *groups* of images requires a fully groupwise objective function. Statistical analysis of the spatial and pixel-value deformations across the set (as defined by the found correspondence), means that these deformations have to be defined with respect to a common spatial and pixel value reference.
We show how the optimal groupwise correspondence can be defined using the Minimum Description Length (MDL) principle, where the definition of the spatial and pixel-value reference is also part of the optimisation. We demonstrate the use of such an objective function for the non-rigid registration of a set of 2D T1-weighted images of the human brain. As regards constructing the optimal reference image, we show that even in the case when substantial portions of the images are missing, the algorithm not only converges to the correct solution, but also allows meaningful integration of image data across the training set, allowing the original image to be reconstructed as the reference image.

1 Introduction

The aim of a non-rigid registration algorithm, as applied to a set of medical images, is to find a 'meaningful' dense correspondence across the whole set of images. There are many methods of *pairwise* registration available (for a review, see [16]), and such registration methods are obviously sufficient for applications such as comparison to an atlas [2]. However, in any application where the statistical analysis of the resulting deformation fields is required – such as the modelling of biological variability, or of assisting in disease diagnosis across the population – performing repeated pairwise registrations over the set of images is, at best, naïve. To facilitate useful statistical analysis, the registration of the group of images needs to be considered as a single problem, so that the parameters of the warps on all of the images lie in a common manifold. We have previously [7] considered a method of non-rigid registration that ensures that there is a common set of knotpoints that define the warps across all of the images. In this paper, we extend that work by considering a groupwise objective function for non-rigid registration.
There is an important distinction to be made between **intra**-subject as opposed to **inter**-subject registration. In intra-subject registration there is often some actual physical process determining the observed deformation (e.g., tissue deformation due to patient position, the needle insertion or organ motion). Alternatively, the deformation may be viewed as the result of some long-term biological process (e.g., natural growth, tumour growth, or atrophy, as in dementia). The most suitable choice of registration algorithm is hence one that closely models the underlying process, leading to physically-based registration algorithms (e.g., [4, 5]), or physically-based models (e.g., [9]) that can be used to evaluate the results of non-rigid registration algorithms. However, in inter-subject registration there is no longer a direct underlying physical or biological process that generates the observed data. We therefore contend that, in the absence of expert anatomical knowledge (i.e., for the case of purely *automatic* registration), the meaning of correspondences should be derived purely from the available data (i.e., the set of images). Further, any statistical inferences that we make about the data should not depend on hypothetical data-generating processes; an assumption that underlies parameter estimation techniques such as maximum likelihood. The Minimum Description Length (MDL) [8] and Minimum Message Length (MML) [15] principles are closely related approaches [1] to model-selection and statistical inference that satisfy these restrictions.
The MDL principle has previously been shown to give excellent results when applied to the correspondence problem in shape modelling [3]. However, naïve attempts at extending the methods described there to images have not been successful [11]. This paper describes the application of the ideas developed in [13, 14].

2 The MDL Principle applied to Groupwise Registration: A Brief Overview

We give here a brief overview of the MDL principle as applied to image registration; for further details, see [13,14]. The key idea is to consider transmitting the full set of quantized images to a receiver, where this image set has been encoded using some type of model. For the case of non-rigid registration, this transmission can be taken to consist of the following parts:

*Joint first authors. Contact: **carole.twining@man.ac.uk, S.R.Marsland@massey.ac.nz**

- The reference example or template – *this defines the common spatial and pixel-value reference for the set.*
- A set of spatial and pixel/voxel value deformations. These can be separated into:
  - A explicitly modelled part – *model parameters and the data as represented by the model.*
  - Any residual deformations.

The model can be either an explicit parametric statistical model (such as a multivariate Gaussian), or the empirical distribution of the actual data (that is, the histogram of the data values). In either case, the receiver reconstructs the reference image, and then applies the specified spatial and pixel-value deformations to this reference, so as to reconstruct *exactly* each image in the training set.
The total transmission/description length $\mathfrak{L}$ can then be written as a sum of corresponding terms thus:

$$\mathfrak{L} = \mathfrak{L}_{\text{reference}} + (\mathfrak{L}_{\text{params:model}} + \mathfrak{L}_{\text{data:model}}) + \mathfrak{L}_{\text{residual}}. \tag{1}$$

2.1 Computing Description Lengths

The actual description lengths for the transmission of one parameter or one piece of data are computed using the fundamental result of Shannon [10] – if there are a set of possible, discrete events $\{i\}$ with associated model probabilities $\{p_i\}$, then the optimum code length required to transmit the occurrence of event i is given by:

$$\mathfrak{L}_i = -\ln p_i \text{ nats}, \tag{2}$$

where the *nat* is the analogous unit to the *bit*, but using a base of e rather than base 2, so that e bits $\equiv 1$ nat. So, for the case of a statistical model, the probability is that given by the model.
The other case we consider is where we wish to transmit an unbounded, quantized data value, or an integer. The two are equivalent, as a quantized data value can always be reduced to an integer. The approximate description lengths are as follows:

$$\textbf{Unsigned Integer: } n \in \mathbb{Z}^+,\ \mathfrak{L}_{\mathbb{Z}^+}(n) \approx \frac{1}{e} + \ln(n) \text{ nats}, \quad \textbf{Signed Integer: } n \in \mathbb{Z},\ \mathfrak{L}_{\mathbb{Z}}(n) \approx \frac{2}{e} + \ln(n) \text{ nats}. \tag{3}$$

We can hence compute the description length for transmitting a quantized, pixellated grayscale image I with N_I pixels according to the image histogram of that image. Suppose that the pixel-values $\{I(A) : A = 1, \ldots N_I\}$ are integers in the range $[1 : N]$, and that there are N_m pixels in the image with the value m, with occupied bins situated at positions $\{m_\alpha\}$. Using this image histogram as the model, this gives the associated probability $p(m) = N_m/N_I$. The transmission then consists of the positions of the occupied bins (assuming a flat distribution over the allowed range), the occupation of each bin, and then finally the ordered set of actual pixel values in the image, encoded using the histogram as model. The description length is hence:

$$\mathfrak{L}_{\text{histogram}} = -\sum_\alpha \ln\left(\frac{m_\alpha}{N}\right) + \sum_\alpha \mathfrak{L}_{\mathbb{Z}^+}(N_{m_\alpha}) - \sum_{A=1}^{N_I} \ln p(I(A)), \tag{4}$$

which is a form of image encoding that we will use later on.

2.2 A Simple MDL Algorithm for Image Registration

We here describe a simple MDL algorithm for image registration. This algorithm has the advantage that it can bootstrap itself from the assumption of the identity transform between image frames, and hence can be used to initialize other more complicated algorithms. We take a set of training images $I_1, \ldots I_{n_s}$, and a reference image I_{ref}, where for each training image we have the spatial transformation t_i between the reference and image planes, as shown in Figure 1. A set of such transformations $\{t_i\}$ is sufficient to define a consistent dense correspondence across the set of images, and in this formulation, these transformations are the only free parameters of the encoding. For a given set of transformations, the reference image is taken as the average of the set of training images, pulled-back into the frame of the reference. We also transmit the discrepancy images given by calculating the discrepancies between each training image and the reference image pushed-forward into the frame of each training image. The description length for transmitting the set of transformations $\{t_I\}$, the reference image, and the set of discrepancy images is then computed, and used as the objective function for defining the optimal set of such transformations. See [13] for further details.

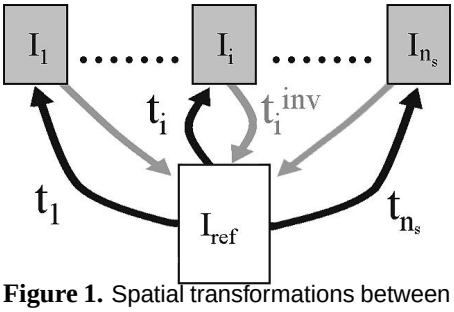


Figure 1. Spatial transformations between reference and image frames.

3 Experiments

3.1 Non-Rigid Registration

As an example, we take as our training set a set of $n_s = 5$ 2D axial T1 MR slices of human brains, which have already been affinely aligned. The images are 8-bit grayscale images of size $N_I = 100 \times 100$. We take as our

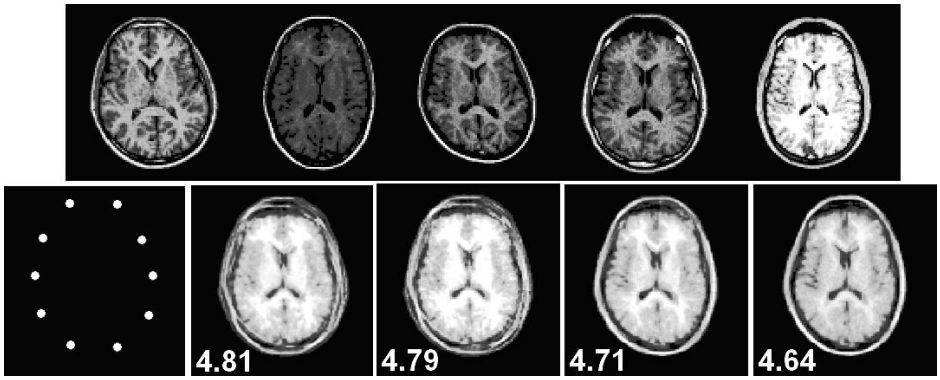


Figure 2. Top: The training set. **Bottom:** The fixed knotpoint positions in the reference frame, and the mean of the aligned images after 0, 2, 6 and 10 iterations, with the transmission length per pixel shown.

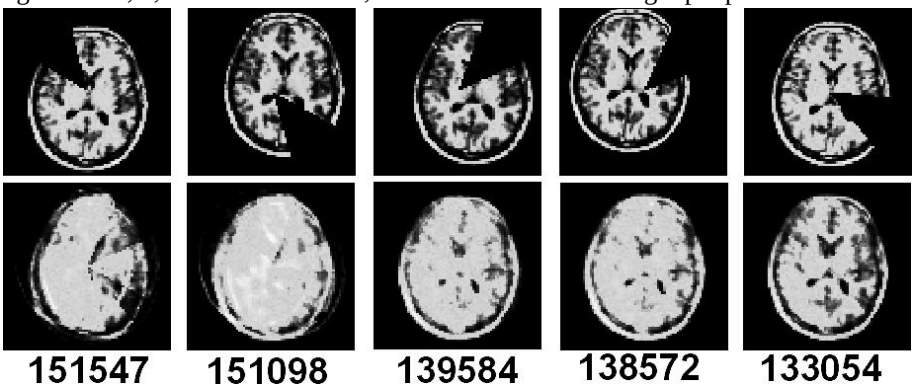


Figure 3. Top Row: The set of training images. **Bottom Row:** The reference image as the registration progresses, with the value of the objective function (the *total* description length for the set in nats).

parameterised set of transformations the biharmonic Clamped-Plate splines (CPS) [12]. The CPS interpolates the motion of a set of knotpoints, hence the parameters of a transformation t_i are the initial and final positions of those knotpoints. The bounding ball for these splines was taken to be the circumcircle of the images. The reference image and discrepancy images are all encoded using the histogram encoding given earlier (4) with $N = 256$ for the reference, and $N = 512$ for the discrepancy images. Following [6], we first generate a set of $n_k = 10$ equiangularly spaced knotpoints around the skull for each image, these being averaged to give the reference image knotpoint positions that remain fixed, and provide us with our spatial reference. For the purposes of illustration, the image knotpoint positions were initialised to the reference knotpoint positions, so that the transformation starts at the identity. Optimising the set of transformations $\{t_i\}$ then corresponds to optimising the set of knotpoint positions on each image. It can be seen in Figure 2 that, as the optimisation proceeds, the reference image sharpens. We see that the structures in the vicinity of the knotpoints are aligned, giving a clear distinction in the reference image between skull, CSF, and the brain surface. The brain structures far from the knotpoints (i.e., the ventricles and sulci) are only approximately aligned, as we would expect. Note also that the final reference does not have the same skull shape as any of the originals. In these results we have only shown the first stage in the registration – as in [6], the registration would be refined by adding more knotpoints, and then re-optimising.

3.2 Optimising the Reference Image

We could have used one of the training examples itself as the reference image – however it is well known that changing the choice of reference can greatly change the final results when it comes to atlas construction. Bhatia et al. [2] perform groupwise registration to a varying spatial reference, yet use a fixed example from the training set as the intensity reference. The problem with such a fixed choice of intensity reference is illustrated in the following example. We take a seed image of a brain slice, and generate a training set of transformed versions of this seed image by translating and resampling. We then obscure part of the brain in each training example, as is shown in Figure 3. It is obvious that using any of these training examples as the intensity reference (e.g., as in [2]) will give poor results, since none of the training examples contain all of the structures present in the seed image. However, as can be seen from the Fig-

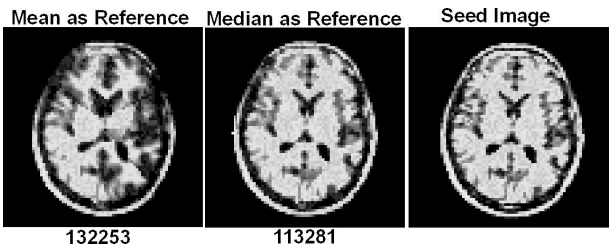


Figure 4.

ure, aligning to the continually-updated mean produces good results, with all the examples being brought into the correct relative alignment. Note, however, that the MDL formulation is not limited just to the choice of the mean of the aligned images as the intensity reference – the values of the reference image are a part of the model, and so could theoretically be optimised over. This is illustrated in Figure 4, where we take the set of transformations given in the previous Figure, but rather than computing the mean, we instead compute the median of the aligned training examples. As can be seen, this not only gives a much smaller description length, but also a reference image that is closer to the seed image.

4 Discussion & Conclusions

We have considered the problem of simultaneously registering entire sets of images, as is necessary for the statistical analysis of image warps to assist in disease diagnosis. In this paper, we have demonstrated that an objective function based on the Minimum Description Length (MDL) principle can be used for this groupwise registration. While here we have demonstrated results using only T1-weighted MRI images, we have shown in previous work [14] that all of the common objective functions used for image registration can be described as modelling choices within the MDL framework, so that the extension to multi-modal images involves merely a change of modelling choice. Similarly, the extension to non-scalar valued images (such as DT-MRI) is also possible, and is currently under investigation.

In the experiments presented in this paper, the reference image was chosen to be the mean or median average image of the aligned training set. We could also have refined this reference image using the MDL objective function, which may have further improved the results – this will be investigated in future work. The experiments that we have presented clearly demonstrate the power of the method: even when different regions of the images are masked off, the algorithm still converges to the correct answer, since there is sufficient information in the entire set of images. This is only possible because the reference image for both spatial and intensity information is a function of *all* of the images in the group. This paper has demonstrated a successful proof-of-concept for the groupwise objective function that we propose. Demonstrating the method on multi-modal images, and in 3D, does not provide any theoretical difficulties, and will be followed-up in the future.

Acknowledgements. Our thanks to A. Fitzgibbon for suggesting the missing data experiment. This research was supported by the MIAS IRC project, EPSRC grant number GR/N14248/01.

References

1. R.A. Baxter and J.J. Oliver. MDL and MML: Similarities and differences. Technical report TR 207, Department of Computer Science, Monash University, Australia, 1994.
2. K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. D. Edwards, and D. Rueckert. Consistent groupwise non-rigid registration for atlas construction. *Proceedings of ISBI 2004*, pages 908–911, 2004.
3. Rh. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor. 3D statistical shape models using direct optimisation of description length. *LNCS*, 2352:3–20, 2002.
4. M. Ferrant, S. K. Warfield, C. R. G. Guttmann, R. V. Mulkern, F. A. Jolesz, and R. Kikinis. 3D image matching using a finite element based elastic deformation model. *Lecture Notes in Computer Science*, 1679:202–209, 1999.
5. A. Hagemann, K. Rohr, H. S. Stiehl, U. Spetzger, and J. M. Gilsbach. Biomechanical modelling of the human head for physically based, nonrigid registration. *IEEE Transactions on Medical Imaging*, 18(10):875–884, 1999.
6. S. Marsland and C. J. Twining. Constructing data-driven optimal representations for iterative pairwise non-rigid registration. *Lecture Notes in Computer Science*, 2717:50–60, 2003.
7. S. Marsland, C. J. Twining, and C. J. Taylor. Groupwise non-rigid registration using polyharmonic clamped-plate splines. *Lecture Notes in Computer Science*, 2879:771–779, 2003.
8. J. Rissanen. *Stochastic Complexity in Statistical Inquiry*. World Scientific Press, 1989.
9. J. A. Schnabel, C. Tanner, A. Castellano Smith, M. O. Leach, C. Hayes, A. Degenhard, R. Hose, D. L. G. Hill, and D. J. Hawkes. Validation of non-rigid registration using finite element methods. *Lecture Notes in Computer Science*, 2082:344–357, 2001.
10. C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423,623–656, 1948.
11. H.H. Thodberg. Minimum description length shape and appearance models. *Lecture Notes in Computer Science*, 2732, 2003.
12. C. J. Twining, S. Marsland, and C. J. Taylor. Measuring geodesic distances on the space of bounded diffeomorphisms. *Proceedings of the British Machine Vision Conference (BMVC)*, 2:847–856, 2002.
13. C. J. Twining, S. Marsland, and C. J. Taylor. Groupwise non-rigid registration: The minimum desription length approach. In *Proceedings of the British Machine Vision Conference (BMVC)*, 2004.
14. C.J. Twining, S. Marsland, and C.J. Taylor. A unified information-theoretic approach to the correspondence problem in image registration. In *Proceedings of International Conference on Pattern Recognition (ICPR)*, 2004.
15. C. S. Wallace and P. R. Freeman. Estimation and inference by compact coding. *Journal of the Royal Statistical Society B*, 54(3):240–265, 1987.
16. Barbara Zitová and Jan Flusser. Image registration methods: A survey. *Image and Vision Computing*, 21:977 – 1000, 2003.

Probabilistic Shape Analysis

Jeong-Gyoo Kim[a*] and J. Alison Noble[a]

[a]Medical Vision Laboratory, Department of Engineering. Science, Oxford University, UK.

Abstract. A shape is modelled as an average planar shape and analysed in terms of a score defined in the model. While most conventional methods of shape analysis have focused on shapes characterised by clear landmarks, in this paper the proposed model is for shapes having few or no clear landmarks. Building on our previous work for generating an average shape, the proposed model additionally provides a concept of a score by which shapes are transformed into numeric variables. This offers a simple and universal method of analysis. It also employs an approximation by basis using Berthilsson's algorithm, which shows a good quality in aligning. The model is applied to an example and its result is compared using a noble score defined in terms of the model to an average generated with a uniform distribution.

1 Introduction

A number of methods in shape modelling have been proposed over the last two decades. Kendall [7] grouped shapes with an equivalence relation and defined shape on a differential manifold, a sphere with a Procrustean metric on it. Bookstein's triangular shape [2] is modelled on a sphere via the complex plane, but with a different metric from Kendall's. Pennec et al [9] modelled a shape as a combination of a feature (such as a point or curve) and a transformation. Both the feature set and transformation set constitute differential manifolds.

Deformable models defined by an energy minimisation mechanism [6] [3] have been an active area of medical imaging and shape analysis research. Bookstein's [3] decomposition of deformation, by affine and non-affine transformations, accelerated research of related topics. A displacement between two sets of landmarks is expressed using the fundamental solution of a biharmonic equation. He adopts bending energy, a bit differently from [6] and formulates a warp function. The whole warp of the displacement is visualised as a thin-plate spline. This method has been widely applied in many areas.

Cootes et al's model for an average shape [4], a Point Distribution Model is efficient and easy to apply and test. In particular, shape variations are described by an eigenstructure in a comprehensive manner. Their model called the Active Shape Model is one of the most commonly adapted methods. However, the intrinsic linearity of the model sometimes results in an average shape that deviates from a population where samples have few clear landmarks.

Sparr [10] models shapes, such as polyhedrons, with an equivalence relation in a projective space. A shape is represented by an annihilator-like concept called Affine Shape; it is an orthogonal complement of a span of bases. Berthilsson and Åström [1] extend Sparr's idea to a shape represented by a continuous curve as a member of the Hilbert space $L^2[0,1]$, the set of square integrable functions on the interval [0,1], and suggest an algorithm for generating bases for the Hilbert space. Combining the ideas of Affine Shape for an infinite dimensional linear space [1] and Active Shapes [4], Ericsson and Åström propose an Affine Invariant Active Shape and show a model generated with their algorithm turns out to be more efficient than a method using Minimum Description Length [5].

Conventional methods of shape analysis mostly characterise a model in terms of landmarks, usually a small number of landmarks and so correspondence among landmarks is of great importance in their models [7], [2], [4]. The manual landmarking process for these methods is tedious and the automated processes currently being developed often has limited accuracy. This paper proposes a method of generating an average of shapes and a definition of a score characterising shape as numeric variables for an analysis of the average. The model provides a way for shape analysis of medical objects having few or no clear landmarks and consisting of dense point-sets. It uses the definition of an average shape suggested previously in [8] where a planar shape is regarded as a continuous function and shapes represented by closed curves are modelled on the Wiener measure space. This paper deals with shapes represented by open curves and adds a concept of a *score* to our previous work. The model for open curves includes parameterisation whilst the previous work [8] does not. For an alignment at the initial stage of modelling, shapes are approximated by a finite number of bases of the Hilbert space $L^2[0,1]$. The bases of the Hilbert space are generated principally with the algorithm for Affine Shape in [1].

The background used in the model and modelling shape represented by open curves are introduced in section 2. The resulting average shape of an example generated from the method and an analysis by its scores are presented in section 3. Section 4 presents a discussion of the proposed method and conclusions of the work.

*jgkim@robots.ox.ac.uk.

2 Background and modelling

Before outlining the contribution of this paper, two core ideas used in modelling are introduced: generating an average shape in [8] and a definition of a *score* to globally characterise shape as a numerical variable.

2.1 Average curve defined on the Wiener measure space

In the authors' previous work [8], a set of planar shapes, represented by closed curves, is considered as a set of continuous functions defined on a bounded interval. The set is regarded as a subset of the Wiener measure space $C_0[a,b]$, the set of all continuous functions defined on the interval $[a,b]$ that vanish at a, and an average function of a given set is constructed via the Wiener measure (a probability distribution). Measure theory provides a generalisation of the concepts of size to arbitrary sets and provides the basic framework for probability theory. This subsection is mainly a summary of the work of [8].

Let us assume we are given a set of m curves (point-sets) representing data, $\{x^1,\dots,x^m\}$, where $x^i \in C_0[a,b]$ for $i=1,\dots,m$. Define a function on the set of m curves by $X_t(x) := x(t)$ for $t \in (a,b]$. It is known that the function X_t is measurable on the Wiener space and follows a Gaussian distribution, $N(0,t)$ [11]. Let $T = \{t_1,\dots,t_n\}$ be a linearly ordered subset of $(a,b]$ and $(X_{t_1},\dots,X_{t_n})$ be a random vector (a vecotr of measurable functions) consisting of them. The setting of a random vector is a difference from [8]. The Wiener integral of the function X_{t_j} for an average is defined by cylinder sets. A cylinder set at each point $t_j \in (a,b]$ is defined with boundaries. The boundaries, α_j and β_j, of cylinder sets are determined by the set $\{x^1(t_j),\dots,x^m(t_j)\}$ as the range of deformation, say, min and max of the set. Hence, the deformation over all x's at t_j is quantified by α_j and β_j. The cylinder sets are formulated by the sets $\mathcal{I}_j = \{x \in C_0[a,b] : \alpha_j < x(t_j) \le \beta_j\}$ for every $t_j \in T$ and depicted in Fig. 1(a). Then the average value $\bar{X}_{t_j}$ of the measurable function X_{t_j} is evaluated for each t_j with the Wiener integral over the cylinder set $\mathcal{I}_j$ and is converted to the Riemann integral in equation (1) [8]:

$$\bar{X}_{t_j} = \frac{\int_{\alpha_j}^{\beta_j} \frac{1}{\sqrt{2\pi(t_j-t_0)}}\, u\, e^{\frac{-u^2}{2(t_j-\alpha)}}\, du}{\int_{\alpha_j}^{\beta_j} \frac{1}{\sqrt{2\pi(t_j-t_0)}}\, e^{\frac{-u^2}{2(t_j-\alpha)}}\, du}. \tag{1}$$

The value $\bar{X}_{t_j}$ is mapped for each t_j. The vector composed of these values is called an average curve of the set of m curves. Notate the vector $\bar{X} = (\bar{X}_{t_1},\dots,\bar{X}_{t_n})$. This average curve $\bar{X}$ always exists within the range of deformation presented in the sample set. In [8], the bounded interval $[0,2\pi]$ is taken for an application to a set of closed curves and the average curve created with the Wiener measure explains deformation better than the average with uniform distribution.

2.2 Score characterising to a numeric variable

Let $\bar{X}$ be the average curve of the given sample of curves $\{x^1,\dots,x^m\}$ created by the method described in subsection 2.1. Let us define a *quasi-variance* of curves in the sample from the average by

$$\sigma^2 := \frac{\sum_{i=1}^m \|x^i - \bar{X}\|^2}{m-1}, \tag{2}$$

where, $\|\cdot\|$ is L^2-norm. The formula (2) is chosen on the ground that a variance is usually described by an average distance between numeric variables in a sample and their mean. With the average curve $\bar{X}$ and the quasi-variance σ^2 in the formula (2), let us define a *quasi-score* of x in a sample by

$$z := \frac{\|x - \bar{X}\|}{\sigma}, \tag{3}$$

where σ is the positive square root of σ^2. The idea of formula (3) is motivated by the fact that normal distributions can be transformed to standard normal distributions by the formula $z = \frac{x-\mu}{\sigma}$, where numerical variables μ and σ are a mean and a standard deviation of the original normal distribution, respectively. Then a shape (continuous function x) is summarised by a score-like single quantity z. The sample of continuous curves $\{x^1,\dots,x^m\}$ are transformed to m numeric variables $\{z^1,\dots,z^m\}$. Hence, the transformation simplifies global information of members of a sample set. The bigger the quasi-score of a shape, the more global the deformation present in the shape. These concepts can be used for cross-analysis between models employing non-numeric variables and/or employing different distributions.

2.3 Modelling

Let us assume that a planar shape is represented by an open continuous curve. Initially each of its coordinates is arc length parameterised and is represented by a pair of real-valued continuous functions $(x(\cdot), y(\cdot))$, where $x, y : [0,1] \to R$. Let $\mathcal{S}$ be the set of those continuous functions x and y. An estimation of a 4-D basis of the Hilbert space $L^2[0,1]$ is generated principally following the algorithm suggested in [1], where reparameterisation is employed to minimise the L^2-norm of an annihilator of the basis. All functions in $\mathcal{S}$ are approximated by the basis: $x \approx \sum_{k=1}^{4} < b_k, x > b_k$, where b_k is a member of the basis and $< \cdot, \cdot >$ is the inner product in the Hilbert space $L^2[0,1]$. Then they are affine transformed and notated by $\mathcal{A}$, a sample set to be modelled. The members of the set of approximated-by-bases and then affine-transformed-curves, $\mathcal{A}$, are again notated by $(x(\cdot), y(\cdot))$.

Once curves have been approximated by the basis and affine aligned, an average curve of $\mathcal{A}$ is fundamentally defined in a similar way to [8]; the set $C_0[0,1]$ is a subset of the Hilbert space $L^2[0,1]$. Let us assume that $\mathcal{A}$ consists of m curves (point-sets), say $\{(x^1, y^1), \ldots, (x^m, y^m)\}$. Two sets of measurable functions $X_t(x) = x(t)$ and $Y_t(x) = y(t)$ for $t \in (0,1]$ are used and their averages are defined as in subsection 2.1. Let $T = \{t_1, \ldots, t_n\}$ be the parameter set, a linearly ordered subset of (0,1]. With the linearly ordered set T, x and y-coordinates are modelled with two random vectors $(X_{t_1}, \ldots, X_{t_n})$ and $(Y_{t_1}, \ldots, Y_{t_n})$, respectively. Then the average value $\bar{X}_{t_j}$ of each component of the random vector $(X_{t_1}, \ldots, X_{t_n})$ is evaluated by using equation (1). The average (as a vector) obtained for x-coordinate is $(\bar{X}_{t_1}, \ldots, \bar{X}_{t_n})$. In the same manner, the average $(\bar{Y}_{t_1}, \ldots, \bar{Y}_{t_n})$ for y-coordinate is obtained. The set of pairs $\{(\bar{X}_{t_j}, \bar{Y}_{t_j}) : j = 1, \ldots, n\}$ defines an average curve of the set $\mathcal{A}$.

A number of methods in shape analysis use a point distribution, mostly a uniform distribution which may well explain small deformations. These are relevant for models with a small number of clear landmarks. However, the distribution does not explain shape variation well where large deformation occurs [8] and may produce an average deviated from populations for a model having few or no clear landmarks [4]. The approach in this paper uses a point distribution of a sample set but differs from conventional methods in that it accommodates a distribution of a continuum simultaneously.

The quasi-score of the sample is evaluated by $z^i = \|(x^i, y^i) - (\bar{X}, \bar{Y})\|/\sigma$, where $\sigma^2 = \sum_{i=1}^{m} \|(x^i, y^i) - (\bar{X}, \bar{Y})\|^2/(m-1)$ as in formulas (3) and (2), respectively. Hence, the non-numeric variable (x^i, y^i) is transformed to a numeric variable z^i and the transformation simplifies global information of members of the sample set. The average can be analysed by measuring the spreadness of the numeric variables $\{z^i\}$. It also can provide a way for further analysis of the sample.

3 Application

The model is applied to data consisting of 32 femurs[1]. The femurs data have various lengths from 89 to 217 (their average length is 131.6875) and are initially parameterised with 130 equally spaced arc length to be as dense as data. The femurs are approximated by the 4-D basis for the Hilbert space as described in subsection 2.3. The mean of relative errors of 32 curves caused by approximation is 0.002592; the relative error is evaluated with the formula $\|(\bar{X}, \bar{Y}) - (x^j, y^j)\|/\|(\bar{X}, \bar{Y})\|$. The approximated curves are affine transformed and the transformed curves are depicted in fig 1(b); affine transformed arc length parameterised curves without approximation are presented in Fig 2(a) for comparison. The curves approximated by bases show far better quality in affine alignment.

The affine transformed curves are modelled with a random vector $(X_{t_1}, \ldots, X_{t_n})$ for their x-coordinate, where X_{t_j} is defined in subsection 2.1. Then the average of each component of the vector $(\bar{X}_{t_1}, \ldots, \bar{X}_{t_n})$ is evaluated with equation (1). The graph of the x-coordinate, cylinder sets and average overlaid on them are illustrated in fig 1(a). In parallel, a random vector $(Y_{t_1}, \ldots, Y_{t_n})$ is defined for y-coordinate and its average $(\bar{Y}_{t_1}, \ldots, \bar{Y}_{t_n})$ of each component of the vector is evaluated. Then, the set of pairs $\{(\bar{X}_{t_j}, \bar{Y}_{t_j}) : j = 1, .., n\}$ represents the average curve of the aligned 32 curves. Fig 1(b) shows the average curve overlaid on the affine transformed curves. To compare with an average evaluated with uniform distribution (a conventional way of generating an average of shapes), the concept of a quasi-score defined in formula (3) is used. The quasi-scores of curves with the Wiener measure are depicted in Fig 2(b) and compared to that with the uniform distribution. As supposed from the plot, the variance of the quasi-scores of the Wiener measure is smaller (0.1029) than that of a uniform distribution (0.3114). As a global analysis, the average with the Wiener measure shows a better representative of the sample than a uniform distribution because the smaller variance of quasi-scores implies the sample is less spread from the average. The

[1] Authors thank Dr. H.H. Thodberg at DTU for kindly providing the femur data.

gap between quasi-scores evaluated with the two distributions tends to be bigger for an example of dissimilar shapes. This is because the average with the Wiener measure better explains deformation than that with a uniform distribution.

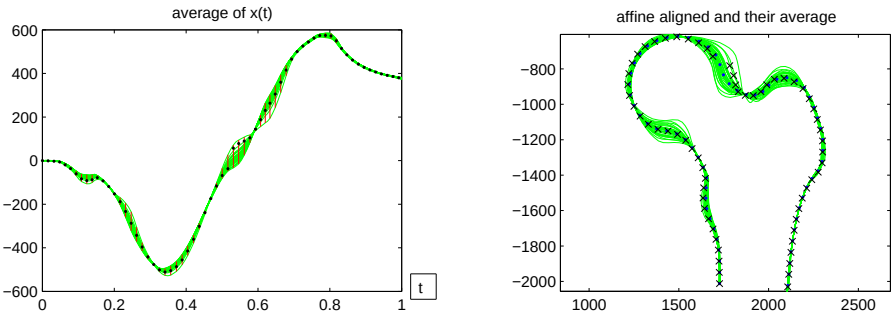


Figure 1. (a) Cylinder sets (depicted by vertical bars) for x-coordinate of the curves in figure (b) and average values overlaid at each t. (b) Average curve with the Wiener measure (plotted with $\times$) overlaid on curves approximated by a basis; average with uniform distribution plotted with $\bullet$.

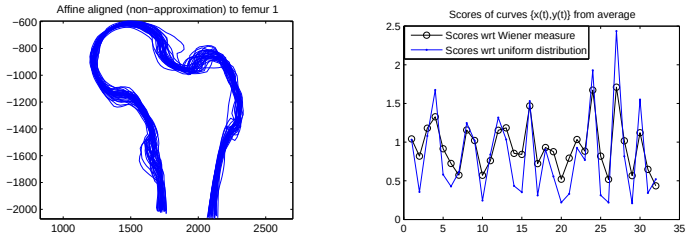


Figure 2. (a) Affine transformed femurs, only arc length parameterised. (b) Quasi-scores of curves on Fig 1(b).

4 Discussion

We have presented a method for generating an average shape represented by open curves and defining a score characterising a shape as a numeric variable. A set of curves are modelled on the Wiener measure space which is defined in terms of Gaussians on a function space. The model explains large deformations which a method employing a uniform distribution usually fail to explain. The proposed definition of a score provides a simple way of analysing shapes in terms of numeric variables. It plays a role of measuring spreadness of a sample with numeric variables in a universal way.

References

1. R. Berthilsson and K. Åström, Extension of affine shape, *JMIV* 11, pp. 119-136, 1999.
2. F.L. Bookstein, Size and shape spaces for landmark data in two dimensions, *Satst. Sci.* 1, pp. 181-221, 238-242, 1986.
3. F.L. Bookstein, Principal warps: Thin-plate splines and the decomposition of deformation, *IEEE Trans. on PAMI* vol. 11:6 pp. 567-585, June 1989.
4. TF. Cootes, CJ. Taylor, DH. Cooper and J. Graham, Active Shape Models - their training and application, *Computer Vision and Image Understanding*, Vol. 61, No. 1 pp. 38-59, 1995.
5. A. Ericsson & K. Åström, An affine invariant deformable shape representation for general curves, *ICCV*, 2003.
6. M. Kass, A. Witkin, D. Terzopoulos, Snakes, Active contour models, *Int. J. Computer Vision*, vol. 1, pp.321-331, 1988.
7. D.G. Kendall, Shape-manifolds, procrustes metrics, and complex projective spaces, *Bull. of London Math. Soc.* 16, pp. 81-121, 1984.
8. J-G. Kim & J.A. Noble, Modelling an average shape *In Proceedings of MIUA*, pp 65-68, BMVA, 2003.
9. X. Pennec, and N. Ayache, Uniform distribution, distance and expectation problems for geometric features processing, *J. of Mathematical Imaging and Vision*, 9, pp. 46-67, 1998.
10. G. Sparr, A common framework for keinetic depth, reconstruction and motion for deformable objects, *Proc. of 4th ECCV*, Cambridge, England, pp. 471-482, 1994.
11. J. Yeh, *Stochastic processes and the Wiener integral*, Marcel Dekker Inc., 1973.

A Generative Statistical Model of Mammographic Appearance

C. J. Rose* and C. J. Taylor

Imaging Science and Biomedical Engineering, University of Manchester, M13 9PT, UK

1 Introduction

We present a generative parametric statistical model of the appearance of entire digital x-ray mammograms. Computer-aided detection (CADe) in mammography has traditionally been treated as a pattern recognition task, where an attempt is made to emulate radiologists' interpretation strategies. We propose instead that CADe be performed via novelty (outlier) detection [1]. Not only are there far more pathology-free than abnormal mammograms from which to learn, but novelty detection would detect *all* abnormal features, rather than specific classes. This requires a model of normal mammographic appearance that allows novel (unlikely) model instances to be identified, suggesting that a statistical model is appropriate. Our model addresses many of the problems associated with modelling the appearance of entire mammograms for novelty detection, but this paper does not focus on using the model in the novelty detection scenario. We propose a method for novelty detection, and offer a discussion of the work, in section 5.

2 Background

Mammograms vary considerably due to differences between women and inconsistencies in the image acquisition process. Women's breasts vary not only in size and shape, but also in their composition. The proportion of glandular to fatty tissue—the density—is variable, with post-menopausal women usually having almost entirely fatty breasts. The number and configuration of ducts varies between women, and the imaging process may capture these in varying degree. Due to the manual placement of the breast in the x-ray equipment, features such as the nipple or pectoral muscle may be absent or only partially visible. A successful model of mammographic appearance must cope with all these sources of variability.

Mammographic appearance has been modelled using physics-based models [2], where the image acquisition process is simulated by computer. However, the synthetic images are not particularly realistic and the approach cannot be used to perform image analysis. Pixel values can be used directly: a parametric model of appearance variation in small image patches is presented in [3], which can be used in both generative and analytical modes, but the explicit assumption of spatial ergodicity does not allow the appearance of the entire breast to be modelled. In [4], wavelet coefficients, computed from small mammographic patches, are statistically modelled using a tree-structured variant of a hidden Markov model. The approach can be used in both generative and analytical modes. Again, the assumption of ergodicity limits the usefulness of the method. Also, due to the structure of the model, the synthetic images have an obvious grid structure. The Active Appearance Model (AAM) [5] is a statistical approach to modelling shape and shape-free appearance. The approach has been applied successfully to a range of computer vision and computer graphics problems. The detailed structure of mammograms is important, but so variable that the AAM approach cannot be applied directly. The approach we have developed combines ideas from the AAM and models of local texture to make the task of modelling mammographic appearance more tractable.

3 Method

We decompose the problem of modelling mammograms by combining an AAM-like model with a spatially ergodic wavelet-based texture model, allowing us to bypass the 'curse of dimensionality'. After outlining a series of preprocessing steps, we present the three components of our model, and show how they can be combined to generate synthetic mammograms.

We assume a training set of mammograms, $\mathcal{B}$, with non-breast regions removed, and normalised such that all breasts 'face' right. As in an AAM, a statistical shape model (SSM) is used to cope with size and shape variation. A set of N 2-D landmark points is required to define the breast borders. These landmarks must correspond across the training set. In the SSM, landmarks are often manually placed and chosen to correspond to intuitive image features (e.g. the corners of the mouth when modelling faces). As mammograms lack reliable features, we seek to automate annotation. A naïve approach is to use landmarks placed at regular intervals on the breast borders,

*christopher.rose@stud.man.ac.uk

starting at a reliable location (e.g. the right-most point on the breast boundary—the approximate location of the nipple). Such landmarks serve as a good first approximation of corresponding points, but the model they produce is not sufficiently specific to be useful. We use a method proposed by Davies *et al.* [6] to improve correspondences across the set of training shapes, yielding an SSM that sucessfully limits illegal shape variation. The resulting Principal Components model has the form:

$$\mathbf{s} = \bar{\mathbf{s}} + P_s \mathbf{b}_s \tag{1}$$

where $\mathbf{s}$ is a shape parameterised by $\mathbf{b}_s$, $\bar{\mathbf{s}}$ is the mean shape, and P_s is a matrix whose columns are a set of eigenvectors of the shape data covariance matrix, sufficient for the model to retain a given proportion of the total variance of the original data. The mean shape $\bar{\mathbf{s}}$ provides a natural canonical reference shape. We warp each segmented breast in $\mathcal{B}$ to the canonical shape, yielding a further set, $\mathcal{N}$, of segmented breasts in a shape-normalised space.

3.1 Modelling the Shape-normalised Appearance

3.1.1 Steerable Pyramid Decomposition

We would like to be able treat the appearance of each mammogram as a point in an appearance space. Although the number of pixels in a mammogram is very large, if there was significant redundancy in the shape-normalised appearance we might still be able to populate the appearance space sufficiently for density estimation to be successful. Unfortunately, this is not the case. To overcome this problem, we use a hierarchical decomposition called the Steerable Pyramid [7]. Images are decomposed in terms of multiple scales and orientations using directional derivative basis functions which range in size and orientation. This allows the coarse and fine structure of the images to be treated separately.

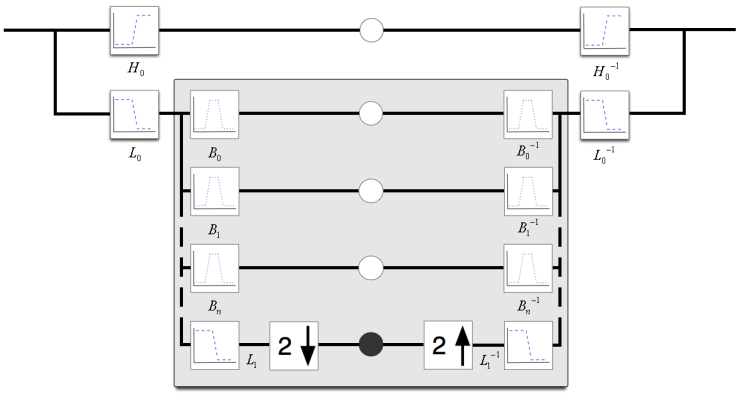


Figure 1. Block diagram for the Steerable Pyramid decomposition. Analysis is shown on the left and synthesis is shown on the right. The dark circle indicates the recursive computation of the shaded region.



Figure 2. The coefficients in the top three levels of a Steerable Pyramid decomposition of a mammogram.

The left-hand side of figure 1 shows the decomposition process (computed using recursive application of oriented bandpass filters to sub-sampled low-pass filtered images) and the right-hand side shows the inverse process of reconstruction. We can think of the decomposition as having a pyramidal structure with discrete levels that correspond to scale, and range from coarse to fine. Each level has a number of oriented sub-bands. Although there are more coefficients in the pyramid than pixels in the original image, the hierarchical structure of the pyramid allows us to decompose our modelling problem further. We can consider the top part of the pyramid (the coarse levels) separately from the bottom part of the pyramid (the fine levels). Figure 2 shows the top three pyramid levels for a mammogram. From $\mathcal{N}$ we form the set of pyramids $\mathcal{P}$.

3.1.2 Approximating Appearance Model

For each pyramid in $\mathcal{P}$, we concatenate the coefficients for the top few pyramid levels into a vector $\mathbf{a}$, which describes the approximate appearance of the shape-normalised mammogram. We again perform Principal Components Analysis (PCA) [1], yielding a model similar to Equation 1:

$$\mathbf{a} = \bar{\mathbf{a}} + P_a \mathbf{b}_a. \tag{2}$$

[1]Initially, the coefficients in each pyramid level are effectively measured on different scales. In this work we often need to use covariance matrices to model the distribution of such data, either for its own sake, or to perform PCA. These techniques work best when we use a common

We then model the distribution of the approximating parameters using a multivariate Gaussian, parameterised by a mean vector $\mathbf{m}_a$ and covariance matrix Σ_a:

$$p(\mathbf{b}_a) \sim N(\mathbf{m}_a, \Sigma_a). \tag{3}$$

To produce a synthetic approximate mammogram (in the normalised shape space), we sample a $\mathbf{b}_a$ from our model of $p(\mathbf{b}_a)$, project back to the natural space via equation 2—to yield the corresponding $\mathbf{a}$ (synthetic coefficients for the top few pyramid levels)—and then reconstruct the pyramid.

3.1.3 Local Textural Detail Model

To model the coefficients in the lower levels—that desribe the fine detail—we assume spatial ergodicity. This makes the problem tractable and is reasonable because we might expect local detail to depend only on tissue type, which is modelled implicitly by the approximating model. A *parent vector*, $\mathbf{b}_t$, which describes local image behaviour, is the set of coefficients on a path through each pyramid level at locations corresponding to a particular pixel in the original image. We seek to model the distribution of parent vectors, $p(\mathbf{b}_t)$, such that coefficients in the detailing levels can be sampled conditionally upon coefficients in the approximating levels. Multivariate Gaussian, or mixture of multivariate Gaussian, representations are ideal for this purpose as there is a closed-form solution for the conditional Gaussian. We use a single multivariate Guassian for computational expediency:

$$p(\mathbf{b}_t) \sim N(\mathbf{m}_t, \Sigma_t). \tag{4}$$

To compute the conditional, we partition the mean vector and covariance matrix of the Gaussian as:

$$\mathbf{m}_t = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix}, \qquad \Sigma_t = \begin{bmatrix} \Sigma_{11} \Sigma_{12} \\ \Sigma_{21} \Sigma_{22} \end{bmatrix} \tag{5}$$

where $\mathbf{m}_1$ corresponds to the unknown dimensions and $\mathbf{m}_2$ corresponds to the known dimensions (the coefficients of the approximating levels); the partitions of the covariance matrix are denoted similarly. If $\mathbf{x}$ is the known part of a particular parent vector, the conditioned mean vector and covariance matrix are computed by:

$$\mathbf{m}' = \mathbf{m}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x} - \mathbf{m}_2), \qquad \Sigma' = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}. \tag{6}$$

To fully populate a pyramid, given coefficients in the approximating part, the conditional is computed and sampled for each parent vector in the pyramid. The fully-populated pyramid can then be reconstructed into the corresponding image in the shape normalised space.

3.2 Joint Shape and Approximate Appearance Model

To complete synthesis we need to be able to warp a synthetic image in the shape-free space to a plausible shape. We need to take into consideration that there may be a relationship between the appearance of a mammogram and its size and shape, for example fatty breasts tend to be large and glandular breasts tend to be small. We model the joint distribution of shape and approximating appearance parameters and condition this model on the approximating parameters to yield a model of plausible shapes for the generated mammogram. Because the number of training examples is likely to be small compared to the dimensionality of this joint space, we use a single multivariate Gaussian:

$$p(\mathbf{b}_s, \mathbf{b}_a) = p(\mathbf{b}_j) \sim N(\mathbf{m}_j, \Sigma_j). \tag{7}$$

Given parameters $\mathbf{b}_a$ describing an approximate mammogram, we can now compute $p(\mathbf{b}_s|\mathbf{b}_a)$ using equation 6 and sample a plausible shape from it.

4 Evaluation and Results

A model was trained with 36 normal mammograms, using 100 boundary landmark points, 7 pyramid levels and 5 orientations. The top three levels were used in the approximating model. The detail model was trained with 100,000 uniformly-sampled parent vectors. It took 24 hours to build the model and 2.5 hours to synthesise a mammogram on a 2.8GHz Intel Xeon processor with 2GB of RAM. Figure 4 shows a real mammogram and three

scale for all dimensions. To achieve this we normalise the data in each dimension either to z-scores, or to a common scale using robust estimation, depending upon the characteristics of the data. For expositional simplicity, the conversion to and from these standard scales is implied in the remainder of the text.

synthetic mammograms generated using the model. Note that our method allows us to produce full-resolution synthetic mammograms. We presented a set of 13 real and 13 synthetic full resolution mammograms to an expert mammography radiologist and a digital mammography computer vision researcher. They gave positive feedback, but could identify the synthetic mammograms by their lack of vasculature, lymph nodes and calcifications. These structures exist at the boundary of the approximating and detailing models, and are not captured by our model.

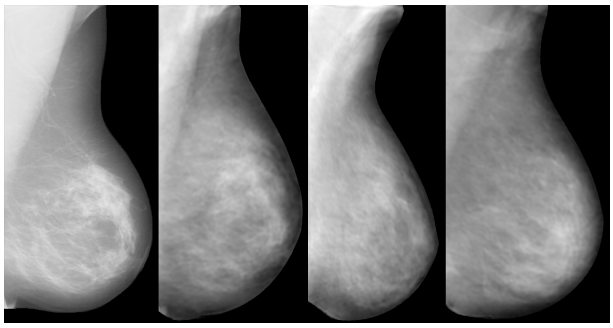


Figure 3. A real mammogram is shown on the left and three synthetic mammograms are shown to its right.

We generated a set of 7 synthetic and 7 real mammograms at 200×140 pixel resolution. We selected real mammograms that lacked strong vascular clues. The synthetic images contained information from both the approximating and detailing models. We recruited five computer vision researchers (though not mammography experts) for a forced choice experiment. After studying a training set of 6 real mammograms, the participants were asked to identify the real mammogram from each of the 49 possible pairings of real and synthetic mammograms. None of the subjects believed that they had been able to identify the real mammograms.

χ^2 analysis showed that two participants did no better than random (one at the 95% level and one at the lenient 99.9% level). The other three participants differed significantly from random, but consistently mistook the synthetic mammograms for the real ones. Between them, the participants correctly identified 75 real mammograms out of 245 (31%). If we allow consistent misclassification to count as correct identification of the real mammograms, the participants collectively identified 191 real mammograms out of 245 (78%). Although the low resolution synthetic images are not always indistinguishable from real mammograms, they are sufficiently convincing to make discrimination difficult. The fact that several subjects consistently mistook the synthetic mammograms for the real ones implies that the differences were very subtle. Although further improvement is required, the results are extremely promising.

5 Discussion

We have developed a generative statistical model of the appearance of entire mammograms. The model combines ideas from the AAM and wavelet-based ergodic texture models to bypass the 'curse of dimensionality'. The model can be used to produce synthetic mammograms, which potentially has pedagogic application. Although our model has only been used for image synthesis so far, it has been constructed to be extended to perform CADe via novelty detection. For example, we could easily look at the likelihood of a local collection of parent vectors under our model of normal appearance. Our synthetic mammograms were evaluated by an expert radiologist and 6 computer vision researchers. Although the model cannot yet capture certain structures found in mammograms, results from psychophysical evaluations of synthetic images generated using our model suggest that the method has significant promise.

References

1. L. Tarassenko, P. Hayton, N. Cerneaz et al. "Novelty Detection for the Identification of Masses in Mammograms." In *4th IEE International Conference on Artificial Neural Networks*, pp. 442–447. Cambridge, UK, 1995.
2. P. R. Bakic, M. Albert, D. Brzakovic et al. "Mammogram synthesis using 3D simulation. I. Breast tissue model and image acquisition simulation." *Medical Physics* **29**, pp. 2131–2139, 2002.
3. C. J. Rose & C. J. Taylor. "A Statistical Model of Texture for Medical Image Synthesis and Analysis." In *Medical Image Understanding and Analysis*, pp. 1–4. 2003.
4. C. Spence, L. Parra & P. Sajda. "Detection, Synthesis and Compression in Mammographic Image Analysis with a Hierarchical Image Probability Model." In L. Staib (editor), *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*. 2001.
5. T. F. Cootes, G. J. Edwards & C. J. Taylor. "Active appearance models." *IEEE Transactions on Pattern Analysis Machine Intelligence* **23(6)**, pp. 681–685, 2001.
6. R. H. Davies, C. J. Twining, T. F. Cootes et al. "A Minimum Description Length Approach to Statistical Shape Modelling." *IEEE Transactions on Medical Imaging* 2002.
7. E. P. Simoncelli & W. T. Freeman. "The Steerable Pyramid: A Flexible Architecture for Multi-Scale Derivative Computation." In *Second International Conference on Image Processing*, volume III, pp. 444–447. IEEE Signal Processing Society, Washington, DC, October 1995.

Wavelet-Features for improved Tumour Detection in DCE-MRI

Birgit Lessmann[a]*, Thorsten Twellmann[b], Andreas Degenhard[a], Tim W Nattkemper[b],
Preminda Kessar[c], Linda Pointon[c], Michael Khazen[c] and Martin O Leach[c]

[a]Theoretical Physics Group, University of Bielefeld, [b]Applied Neuroinformatics Group, University of Bielefeld,
[c]Clinical MR Research Group, Institute of Cancer Research, London

Abstract. In applications for automatic tumour detection using DCE-MR-Images it is often a problem to distinguish the dynamic enhancement of tumour tissue, i.e. the relevant information, from other enhancing tissues. In this work we propose the use of a wavelet-filtering approach to characterise different enhancement behaviour.

1 Introduction

Dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) is currently being investigated for the early detection of breast cancer. The advantage of administering a paramagnetic contrast medium (Gd-DTPA) is twofold. Firstly, suspicious tissue often enhances and thereby yields an improved image contrast. Secondly, classification of lesions assumes a fast uptake of the contrast agent followed by a washout to indicates malignancy, while a slow uptake characterises benign lesions [1]. In the UK Breast Screening Study, premenopausal women at high genetic risk of developing breast cancer, are being evaluated using a dynamical imaging and analysis protocol [2]. In an experimental trial high-sensitivity 3D MR examination are performed. Two images are acquired prior to (pre-contrast) and four or five images are acquired after contrast agent injection (post-contrast), each with an acquisition time of 90 secs.
In this work we propose to investigate the feasibility of automatic tumour detection by ROC analysis utilising subtraction images in a similar way to the radiological approach in the clinical work. In general, sensitivity is affected by additional contrast enhancement within normal breast parenchyma or within the heart. In our approach we increase the sensitivity for high specificity by applying a wavelet filtering method.

2 Methods

2.1 Wavelet-Analysis

Wavelet analysis, including multiresolution analysis, enables to assess the scale-dependent information in signals and images [3]. In this mathematical theory a signal f is decomposed using the *Discrete (Dyadic) Wavelet Transform* into a basis of shifted and dilated versions of a basic wavelet or *mother wavelet* ψ [4]

$$f(x) \;=\; \sum_{(j,k)} d_{j,k}\psi_{j,k}(x)\;, \qquad \text{with} \quad \psi_{j,k}(x) \;=\; 2^{j/2}\psi(2^j x - k)\;. \tag{1}$$

Here the index j indicates the dilation or *scaling step* while k refers to translation or shifting. The wavelet coefficients $d_{j,k}$ are given by the scalar product $d_{j,k} = \langle f(x), \psi_{j,k}(x)\rangle$ or $d_{j,k} = \langle f(x), \tilde{\psi}_{j,k}(x)\rangle$ in case of *biorthogonal wavelets* with the dual wavelet $\tilde{\psi}$ [4]. An efficient calculation of these coefficients is accomplished by the *Fast Wavelet Transform* (FWT), an algorithm allowing the coefficients to be calculated in a stepwise manner. To perform a FWT a scaling function $\phi(x)$ is required such [4]

$$\phi(x) = \sqrt{2}\sum_k h(k)\phi(2x-k) \qquad \text{and} \qquad \psi(x) = \sqrt{2}\sum_k g(k)\phi(2x-k)\;. \tag{2}$$

The coefficients $h(k)$ and $g(k)$ are termed *Filter coefficients* and determine the wavelet. On the first scale the signal is decomposed into its *details* and the remaining signal, i.e. the *approximation*, reflecting the particular scale. The details are described by the wavelet coefficients of this scale while the approximation is represented by scaling coefficients corresponding to the scaling function. The procedure can be iterated by a further decomposition of the approximation into details and approximation of the next coarser scale. A decomposition tree for three decomposition steps is shown in figure 1.

*lessmann@physik.uni-bielefeld.de

2.2 Wavelet-Filtering

In this work a discrete wavelet decomposition for T1-weighted DCE-MR 3D-Images (Volumes) comprising 18 different datasets (patients) is performed. A pre-contrast image together with five (dataset A) or four (dataset B) consecutive post-contrast images are provided to the decomposition scheme. Afterwards a filter in the space of the wavelet coefficients is applied. The filtering method developed detects those time series of coefficients, which are assumed to characterise the tumour [1].
The change in the ratio of signal intensities (SI) between fast enhancing tumour tissue and surrounding tissue with time produces characteristic coefficients on consecutive scales which are included in our filter approach. However, it is observed that the enhancement located in the region of the heart can produce similar characteristics, making a clear separation difficult. Therefore only time series of coefficients, that follow a well-described behaviour predicted for tumour tissue, are retained. This enables tumour enhancement to be distinguished from irregular enhancement of the heart.

2.3 Receiver Operator Characteristic analysis

Receiver Operator Characteristic (ROC) analysis is a useful technique to visualise and measure the performance of classifiers. In our approach subtraction images are computed from the pre-contrast and the first post-contrast volume in the dynamic sequence and are compared to an expert label (a binary mask). All voxels above a specific threshold are classified positively (i.e. as a tumour voxel) and all others negatively. Comparison with the expert label allows the correctly classified voxels (true negative (tn) and true positive (tp)) and the incorrectly classified voxels, i.e. the false negatives (fn) and false positive (fp) voxels, to be calculated for each threshold value. In the ROC graph the true positive rate (sensitivity) (tp rate $= \frac{tp}{tp+fn}$) is plotted against the false positive rate (fp rate $= \frac{fp}{fp+tn}$). The false positive rate is related to the term specificity by specificity $= \frac{tn}{fp+tn} = 1 -$ fp rate [5].

2.4 Machine Learning Approach

We used a *Fisher Linear Discriminant Analysis* (FLDA) [6] to examine the value of our approach as a preprocessing step. Discrimination utilising the FLDA is based on the idea of finding a linear combination $\mathbf{w}$ of the features $\mathbf{x} = (x^{(1)}, \ldots, x^{(n_t)})$ that maximises the ratio between the *between-class scatter* and the *within-class scatter* described by the corresponding matrices $\mathbf{S_b}$ and $\mathbf{S_w}$, respectively. Regarding the binary case, an analytical solution is available which maximises the *Rayleigh coefficient*

$$J(\mathbf{w}) \;=\; \frac{\mathbf{w^T S_b w}}{\mathbf{w^T S_w w}} \qquad \text{using} \qquad \mathbf{w} = \mathbf{S_w^{-1}}(\mathbf{m_1} - \mathbf{m_2}) \tag{3}$$

with class specific mean vectors $\mathbf{m_i}$. The algorithm has no further hyperparameters and is computational inexpensive even for huge data sets. The binary classification for an example $\mathbf{x_j}$ is obtained by $\tilde{y}_j = \text{sgn}[\langle \mathbf{w} \cdot \mathbf{x}_j \rangle + b]$ with a properly chosen bias value b.

2.5 Analysis protocol

After the filtering procedure is completed subtraction images are reconstructed. These filtered subtraction images are compared with the original subtraction images on the basis of ROC analysis.
Two types of filtering methods are distinguished. In the first method (Filter 1) all scales contribute to the filtering procedure whereas in the second method (Filter 2) the filtering is restricted to the first four scales. The latter approach is motivated by the observation that a coefficient describing coarse scale phenomena may contain information of different enhancing structures in the image, making an efficient filtering process difficult.
Three different types of mother wavelets are employed to investigate possible dependencies. These comprise the *Cohen-Daubechies-Feauveau* (CDF)(2,2)-wavelet, the *Haar-wavelet* and the *Daubechies(4)-wavelet* (DB4, sometimes also called DB2). Finally, we tested the filtering method as a preprocessing step for machine learning classification algorithms, examined by performing ROC analysis after classification as shown in figure 2.

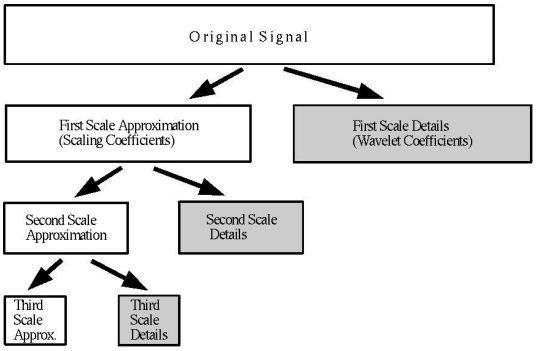


Figure 1. Decomposition tree for three decomposition steps, the length of the bars indicates the number of coefficients

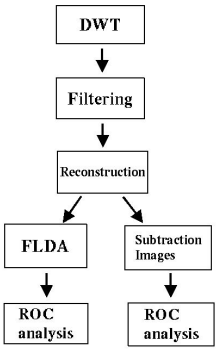


Figure 2. Complete analysis protocol

3 Results

3.1 ROC-Analysis with subtraction images

For dataset A and dataset B the results for both types of filters are compared to the results obtained from the raw data. Tables 1 and 2 show the measured areas under the ROC-Curves (AUC) for high specificity parameters (0.95 - 1) with respect to the different filtering methods. The ROC-Curves are calculated according to the explanations provided in 2.3 and the corresponding AUC values have been normalised to the specific area.

dataset A	malignant cases					benign cases		
case	1	2	3	4	5	6	7	8
raw data	0.82	0.67	0.57	0.81	0.72	0.00	0.79	0.56
filtered data (Filt. 1)	**0.84**	**0.80**	**0.70**	**0.86**	0.69	0.01	0.77	0.55
filtered data (Filt. 2)	**0.91**	**0.81**	**0.58**	**0.87**	0.16	0.00	**0.86**	0.54

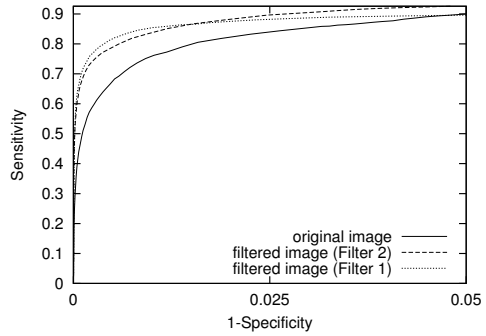
Table 1. Normalised AUC values (dataset A) for specificity ≥ 0.95, values shown bold have increased after filtering

dataset B	malignant cases						benign cases			
case	9	10	11	12	13	14	15	16	17	18
raw data	0.45	0.60	0.65	0.57	0.35	0.30	0.83	0.65	0.33	0.26
filtered data (Filt. 1)	0.30	**0.73**	0.52	0.51	**0.37**	**0.50**	0.60	**0.70**	**0.38**	0.19
filtered data (Filt. 2)	0.41	0.60	**0.71**	0.53	0.27	**0.31**	0.61	**0.74**	0.17	0.12

Table 2. Normalised AUC values (dataset B) for specificity ≥ 0.95, values shown bold have increased after filtering

When applying the CDF(2,2)-wavelet the sensitivity increases considerably for high specificity with respect to the majority of the malignant cases. Case 5, which does not benefit from the different filtering approach, is the image of a patient having an implant with an adjacent tumour. Case 9 also not benefitting of our approach is an image sequence suffering from strong movement of the patient - a problem which cannot be handled by this algorithm.
Figure 3 a) and b) shows as an example the ROC-Curves for the 4th case and the amount of coefficients retained on different scales.
The dataset B have been acquired by using a slightly different imaging protocol, providing only four post-contrast images or in other words, less information. Therefore this dataset produces results with reduced quality as compared to dataset A. This corresponds with the observation, that the amount of coefficients retained for dataset B is remarkably higher than in the cases of dataset A. The best performance for malignant cases has been achieved by applying the first filtering method. Obviously several cases benefit from a less strict filtering (Filter 2).
In contrast to the CDF(2,2)-wavelet the other mother wavelets show to be not capable for this application. In figure 3 c) the AUC values corresponding to different mother wavelets for the 4th case are shown. While the Haar- and the DB4-wavelet lead to poor results, i.e. an decrease in sensitivity, the CDF(2,2)-wavelet, also known as 5/3-Filter, has been shown to be quite capable for these kind of filtering applications. This wavelet is also used in the JPEG2000 compression technique [7].
The proposed filtering method does not succeed when applied to weakly enhancing benign lesions. This is expected, since it is constructed to detect and differentiate strong enhancement. The benign cases 7 and 16 benefitting from our approach are strong enhancing fibroadenomas. In table 3 the results of the FLDA classification

a)



b)

Scale	amount of retained coefficients	
	Filt. 1	Filt. 2
1	0.07	0.07
2	0.06	0.06
3	0.05	0.05
4	0.04	0.04
5	0.02	1
≥ 6	0.01	1

c)

mother wavelet	AUC$\geq$ 0.95	
	Filt. 1	Filt. 2
Haar	0.76	0.81
Daub(4)	0.77	0.78
CDF(2,2)	0.86	0.87

Figure 3. 4th case, a) ROC-Curve for original and filtered image (CDF(2,2)), b) retained coefficients after filtering (CDF(2,2)), c) AUC values for specificity$\geq$0.95 and three different mother wavelets

approach applied to the original data and the filtered data by using Filter 1 are shown. Training with both datasets would require the image sequence of dataset A to be restricted to the pre-contrast and only four-contrast images. Therefore only dataset A is included in the FLDA approach. For the training all datasets are used except the one classified in the next step. This procedure is iterated permuting the classified dataset (*leave-one-out approach*). The results are similar to the subtraction images i.e. the benign cases do not benefit from filtering as opposed to the malignant cases. One exception is case 3, which does not show the expected increase of the AUC value.

dataset A	malignant cases					benign cases		
case	1	2	3	4	5	6	7	8
raw data	0.67	0.59	0.64	0.81	0.55	0.57	0.00	0.69
filtered data (Filt. 1)	**0.73**	**0.68**	0.63	**0.82**	**0.60**	0.37	0.00	0.54

Table 3. Results of FLDA classification - normalised AUC values for specificity $\geq$ 0.95 (dataset A), values shown bold have increased after filtering

4 Summary and Discussion

Our results clearly demonstrate that automatic tumour detection and classification algorithms can benefit from applying multiscale filtering algorithms. The described wavelet-filter-algorithm has clear potential for distinguishing the different types of enhancement. More work is needed to investigate the behaviour of the contrast agent on the coarser scales.

Acknowledgements

Birgit Lessmann acknowledges support by the Federal State Nordrhein-Westfalen and the Graduate programme “structure formation processes”. The authors thank Oliver Schulz (Max-Planck-Institute Dortmund) for the software resources he provided.

References

1. S. H. Heywang-Koebrunner & R. Beck. *Contrast-Enhanced MRI of the Breast.* Springer-Verlag, Berlin, 1996.
2. UK MRI Breast Screening Study Advisory Group. “Magnetic resonance imaging screening in women at genetic risk of breast cancer: imaging and analysis protocol for the UK multicentre study.” *Magn. Reson. Imaging* **18**, pp. 765, 2000.
3. S. Mallat & S. Zhong. “Characterization of signals from multiscale edges.” *IEEE Trans. Patt. Anal. Machine Intell.* **14**, pp. 710, 1992.
4. I. Daubechies. *Ten Lectures on Wavelets*. CBMS-NFS Series Appl. Math., SIAM, 1991.
5. T. Fawcett. “ROC Graphs: Notes and Practical Considerations for Researchers.” *HP Labs Tech Report HPL-2003-4* 2003.
6. R. Duda, P. Heart & D. Stork. *Pattern Classification*. Wiley, New York, 2001.
7. C. Christopoulos, A. Skodras & T. Ebrahimi. “The jpeg2000 still image coding system: an overview.” *IEEE Trans. on Consumer Electronics* **46**, pp. 1103–1127, 2000.

Visualisation of Breast Tumour DCE-MRI Data using LLE

Claudio Varini[a]*, Tim W Nattkemper[a], Andreas Degenhard[a] and Axel Wismueller[b]
[a]Applied Neuroinformatics Group, Faculty of Technology and Physics, University of Bielefeld, Germany

[b]Digital Image Processing Group, Department of Radiology, University of Munich, Germany

Abstract. The usage of dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) for breast cancer examination is currently undergoing intensive research. The data set of each patient is composed of a sequence of three-dimensional images. The Three Time-Point (3TP) method is a well-known model-based technique for visualising the signal dynamics with customised colours. We propose a new approach based on unsupervised machine learning which makes use of the recently proposed Locally Linear Embedding (LLE) algorithm. The first results are in agreement with those obtained by 3TP with the advantage that no model is required.

1 Introduction

The usage of dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) on female breast cancer diagnosis is currently undergoing intensive research. During the examination, the patient is imaged before the injection of a contrast agent and repeatedly thereafter, thereby monitoring the uptake of the contrast agent in the tumour tissue over time. These uptake curves provide valuable information about the vascularity of the tissue and may help to distinguish between benign and malignant lesions [1]. For the purpose of analysing DCE-MRI data, the Three Time-Point (3TP) method is an established, model-based visualisation technique [2]. Here the washin and washout phases are evaluated by analysing the changes of the signal intensity values at three selected time-points, one pre-contrast (t_0) and two post-contrast (t_1, t_2), along the uptake curve of each pixel. The washout is estimated from the values of the two post-contrast intensities $[I(t_1), I(t_2)]$ according to one of the three patterns coded by colour hues: (1) slow washout with $I(t_1) < I(t_2)$ coded blue, indicating benign behaviour; (2) moderate washout with $I(t_1) = I(t_2)$ coded green, suspicious case; (3) fast washout with $I(t_1) > I(t_2)$ coded red, indicating malignant behaviour. The washin is estimated from the first two time-points by $[I(t_1) - I(t_0)]/(t_1 - t_0)$ and its magnitude is coded by colour intensity [3]. By a model-based calibration map, hue and colour intensity of each pixel are then related to two pathophysiological features: microvascular permeability (K), which is equal to the microcapillary surface area times permeability, and the extracellular volume fraction (υ_e), which determines the volume fraction accessible to the contrast agent. The three time-points must be judiciously selected in order to provide sufficient coverage for the temporal characteristics observed [3].
Here we propose a novel and non model-based approach to the visualisation of breast DCE-MRI data based on unsupervised learning by the Locally Linear Embedding (LLE) algorithm, and we compare our technique with the 3TP method. The recently proposed LLE algorithm allows to perform dimensional reduction of nonlinear manifolds [4]. In LLE, each data point is approximated by a linear combination of its n-nearest neighbours. These are usually chosen according to the Euclidean distance metric in the data space. The entire data structure is then mapped into a low-dimensional space and, at the same time, the linear combinations are preserved to the greatest possible extent. We apply LLE to a 6D data set comprising the MR measurements of 12 breast tumours, six benign and six malignant cases. We reduce the data to three dimensions and the coordinates of the data points in the embedding are encoded within the RGB colour scheme to yield a colour image of each tumour, thereby visualising the dynamics of the contrast agent within the tissue by means of customised colours. The absence of a model in our method is compensated by processing 12 tumours data at once by LLE. In this way, one obtains a low-dimensional coordinate system where the data points of a single tumour are localised with respect to the contrast characteristics of the other tumours; such coordinates, in other words, are mapped by learning from the entire data set. In fact, LLE projects multi-dimensional data into a lower-dimensional embedding by preserving the original global structure, i. e. points presenting similarities in the data space lay close to each other in the embedding space. Therefore, time-series having similar behaviour are mapped close to each other in the embedding space, and consequently the respective voxels will have a similar hue.

2 Materials and methods

The data of each patient comprises six three-dimensional images of the full breast taken within an interval of 110 s. The first frame was acquired before the bolus injection of a contrast agent, thereafter followed by the remaining five measurements. The imaging process was performed with a 1,5 T system (Magneton Vision, Siemens, Erlangen,

*Email: varini@physik.uni-bielefeld.de

Germany) equipped with a dedicated surface coil to enable simultaneous imaging of both breasts. First, transversal images were acquired with a STIR (short tau inversion recovery) sequence (TR$=5600$ ms, TE$=60$ ms, FA$=90°$, TI$=150$ ms, matrix size of 256×256 pixels, slice thickness 4mm), then a dynamic T1 weighted gradient echo sequence (3D FLASH) was performed (TR$=12$ ms, TE$=5$ ms, FA$=25°$) in transversal slice orientation with a matrix size of 256×256 pixels and an effective slice thickness of 4 mm. The data set is composed of 12 tumours, six benign and six malignant. The projection in a three-dimensional embedding was performed using LLE with parameters $n = 25$ and $\Delta = 0.0278$. For such values, the LLE output is quite stable, as shown in [5]. The 3D coordinates were then encoded within the RGB colour space. A 2D reduction of the data space was also performed in order to illustrate more clearly the relationship between the 3TP and LLE colour mapping (see Fig. 1). Concerning the 3TP method, the three time-points used were the ones of the pre-contrast, first and last post-contrast image, and they were chosen according to [3].

2.1 The LLE algorithm

The LLE algorithm is based on three steps involving standard methods of linear algebra. Its input consists of N D-dimensional vectors $\mathbf{X}$. In the first step, one identifies n neighbours for each data point $\mathbf{X}_i$. Different criteria for the selection of the neighbours can be adopted; the simplest possibility is to choose the n-nearest neighbours according to the Euclidean distance. By minimizing the following error function (step 2)

$$\Psi(W) = \sum_i |\mathbf{X}_i - \sum_j W_{ij}\mathbf{X}_j|^2 \tag{1}$$

subject to the constraints $\sum_j W_{ij} = 1$ and $W_{ij} = 0$ if $\mathbf{X}_i$ and $\mathbf{X}_j$ are not neighbours, one obtains the linear coefficients that reconstruct each data point from its neighbours. With the above constraints, the quadratic problem can be simplified to a linear system and the weights can be computed in closed form. If the number of neighbours is greater than the input dimension ($n > D$), the solution turns out to be not unique and the linear system must be conditioned with a small term Δ, as described in [4]. The last step of the algorithm consists in mapping each data point $\mathbf{X}_i$ into a low dimensional vector $\mathbf{Y}_i$, such that the following embedding error function results minimised:

$$\Phi(Y) = \sum_i |\mathbf{Y}_i - \sum_j W_{ij}\mathbf{Y}_j|^2 \tag{2}$$

under the condition $\frac{1}{N}\sum_{i=1}^{N} \mathbf{Y}_i\mathbf{Y}_i^T = I$. The weights are kept fixed in order to preserve the local neighbourhood of each data point. The most straightforward method for computing the d-dimensional coordinates, where $d < D$, is to find the bottom $d + 1$ eigenvectors of the sparse matrix

$$M = (I - W)^T(I - W). \tag{3}$$

These eigenvectors are associated with the $d + 1$ smallest eigenvalues of M. The bottom eigenvector, whose eigenvalue is closest to zero, is the unit vector with all equal components and is discarded.

3 Results and discussion

The correlation between the 3TP and LLE colour mapping is shown in Fig. 1. In (a), the 2D LLE embedding is plotted with colours given by the 3TP method. We can observe that most of the time-series are mapped by LLE in agreement with the 3TP analysis. In fact, three different clusters given by red, green and blue can be easily detected, suggesting that there exists a relation between the LLE coordinates and the 3TP colour. Consequently, the LLE coordinates reflect pathophysiological features of the tissue. Points with slow washout are localised in the upper part of Fig. 1(a), while points having faster washout are progressively mapped into the lower region. The marker of each point specifies whether it belong to a benign or a malignant lesion. One can see that all points belonging to benign tumours are clustered by LLE only in the upper part of the figure. In the lower region there are some malignant points with colour blue, i.e. according to the 3TP method, they are labelled as benign. Conversely, the LLE algorithm maps them in the malignant part. By showing the correspondent contrast characteristics, we observe that such points are characterised by an anomalous washout phase. The plots show evidence of a clear washout behaviour, followed however by a signal increase. This was caused by a slight movement of the patient during the MRI acquisition and has not been noticed before. In this case, 3TP fails to interpret these time-series characteristics, confirming a limitation of the method that has already been mentioned in [6]; in contrast, LLE is able to detect such abnormalities by virtue of processing the entire data structure.
In Fig. 1(b) we show the 2D LLE coordinates with the corresponding colours given by the 3D LLE embedding.

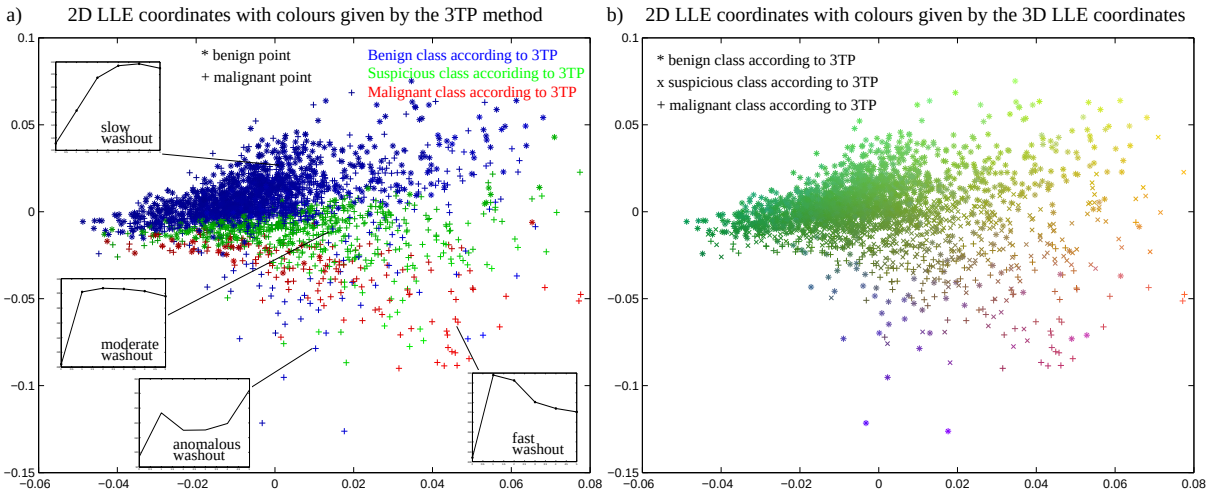


Figure 1. (a) 2D Embedding coordinates from the LLE algorithm. Colours correspond to those obtained by the 3TP method considering the respective time-series in the initial data space. Star-shaped points belong to benign lesions; plus-shaped points belong to malignant lesions. Most of the benign points are localised in the upper part, in agreement with the 3TP analysis (encoded by the blue colour according to 3TP). The majority of the points with moderate washout (suspicious behaviour) are approximately in the central region of the mapping (encoded by the green colour according to 3TP). Most of the points showing fast washout (malignant behaviour) are localised in the lower part of the picture (encoded by the red colour according to 3TP). Some points classified as benign by 3TP are instead mapped close to data points belonging to malignant lesions using LLE . The respective contrast characteristics are also shown and we can note an untypical washout phase due to patient motion. (b) 2D LLE embedding with colours determined by the coordinates of the 3D LLE embedding. The marker of each point gives information about the colour that such point has with respect the 3TP method.

The marker of each point reflects the colour scheme if we would have applied the 3TP method: star-shaped points are blue in 3TP; x-shaped points are green and plus-shaped ones correspond to red points in 3TP. In LLE, points characterised by slow washout (blue points in 3TP) are encoded as green, green-yellow. The green points in the 3TP method are encoded by LLE as dark-green, orange. Most of the points with a fast washout (red colour in 3TP) are then mapped with red and blue hues. In particular, in LLE the blue points are those presenting an anomalous washout phase as described above. The images of the 12 lesions obtained by 3TP and LLE are shown in Fig. 2 (malignant cases) and Fig. 3 (benign cases). Both methods generally show qualitative agreement, and results demonstrate that benign tumours are largely homogeneous, while the malignant ones are more heterogeneous, in agreement with the experience of radiologists [7]. In addition, a strong correlation between the variance of the colours for both methods is observed. We also point out that the images of the first, fifth and sixth malignant lesion, and of the first three benign cases are particularly consistent. By observing the images of the fourth and sixth benign cases in Fig. 3, we note that some voxels are mapped as malignant by 3TP (red hue). The respective images obtained by LLE are in contrast more homogeneous, thereby showing higher specificity. This is particularly valuable, since the DCE-MRI technique is known to present high sensitivity, but a wide range of specificities, with values varying from 30% to 100% [8].

4 Conclusions and Future Work

We presented a novel approach to the visualisation of breast tumours in MR imaging based on unsupervised learning by making use of Locally Linear Embedding (LLE), an algorithm for dimensional data reduction able to unfold nonlinear manifolds. At first, we reduced the 6D data space comprising the MR measurements of 12 breast lesions (six benign and six malignant) to a 3D embedding. We then encoded the new coordinates within the RGB colour space to yield a parametric image of each tumour, with colours reflecting pathophysiology characteristics of the tissue. To test the effectiveness of our approach, we compared it with the well-established Three Time-Point (3TP) technique for tumour visualisation. Our analysis showed sufficient agreement between both methods. In addition, our technique resulted to be more robust with respect to the issue of patient motion during the MRI acquisition. The principal advantages of our method as compared to 3TP are that (1) it processes the entire data and (2) it is model free. Its main shortcoming is that the correlation between hue and tissue pathophysiology is not constant, but depends on the data set provided. Indeed, using other tumour data, the LLE projection could have different coordinates values, and consequently the correlation between characteristics of the tissue and hues

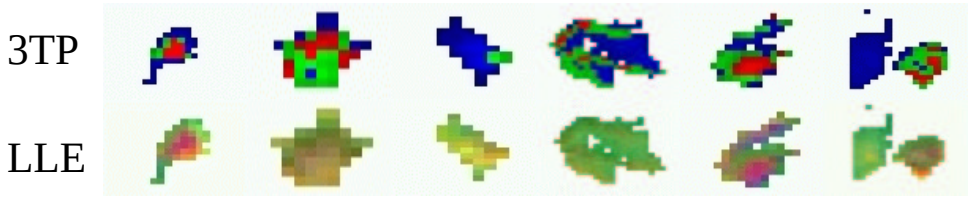


Figure 2. Malignant lesions

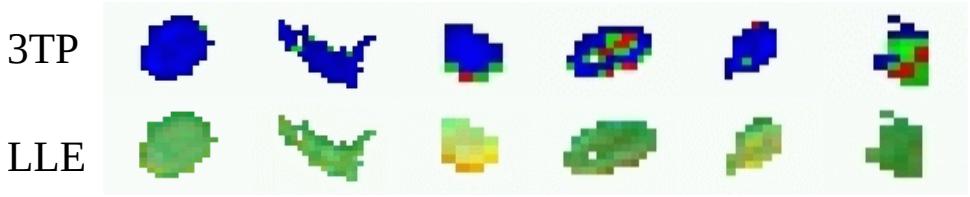


Figure 3. Benign lesions

could result different from what is written above. In the case of a new data set, it is therefore necessary to refind a correlation between dynamics of the contrast characteristics and colour by using a colour map in the same manner as shown in Fig. 1. In future work, a pixel-by-pixel comparison between the LLE and 3TP method will be addressed. Moreover, it may be interesting to apply our technique to tumours in other organs for which a model describing the kinetics of the contrast agent in the tissue does not exist.

5 Acknowledgment

This work was supported by research grants (DFG grant GK-231) received from the *Deutsche Forschungsgemeinschaft* (*German Research Foundation*) within the *Graduiertenkollegs Strukturbildungsprozesse* (*Graduate Program Structure Formation Processes*), University of Bielefeld, Germany. The authors would like to thank the clinical department of the University of Munich, Germany for providing and labelling the data used in this study.

References

1. The UK MRI Breast Screening Study Advisor Group, Imaging Advisors and Staff. "Magnetic Resonance Imaging Screening in Women at genetic Risk of Breast Cancer: Imaging and Analysis Protocol for the UK Multicenter Study." *Magnetic Resonance Imaging* **18**, pp. 765–776, 2000.
2. H Degani, V Gusis, D Weinstein et al. "Mapping Pathophysiological Features of Breast Tumors by MRI at high spatial Resolution." *Nature Medicine* **3(7)**, pp. 780–782, 1997.
3. D Weinstein, S Strano, P Cohen et. al. "Breast Fibroadenoma: Mapping of pathophysiologic Features with three-time-point, contrast enhanced MR imaging-Pilot study." *Radiology* **210**, pp. 233–240, 1999.
4. S T Roweis and L K Saul. "Nonlinear Dimensionality Reduction by Locally Linear Embedding." *Science* **290**, pp. 2323–2326, 2000.
5. C Varini, T W Nattkemper, A Degenhard et al. "Breast MRI Data Analysis by LLE." In *Proc. of Int. Joint Conf. on Neural Networks*. Budapest, Hungary, Jul 2004.
6. F Kelcz, E Furman-Haran, D Grobgeld et al. "Clinical Testing of High-Spatial-Resolution Parametric Contrast Enhanced MR Imaging of the Breast." *AJR* **179**, pp. 1485–1492, 2002.
7. E Furman-Haran, D Grobgeld, F Kelcz et al. "Critical Role of Spatial Resolution in Dynamic Contrast-Enhanced Breast MRI." *Journal of Magnetic Resonance Imaging* **13**, pp. 862–867, 2001.
8. M Su, Y Cheung, J Fruehauf et al. "Correlation of Dynamic Contrast Enhancement MRI Parameters with Microvessel Density and VEGF for Assessment of Angiogenesis in Breast Cancer." *Journal of Magnetic Resonance Imaging* **18**, pp. 467–477, 2003.

Automatic Capillary Measurement and Classification

Kaiyan Feng[a] *, P. D. Allen[a], T. Moore[b], A. L. Herrick[b], and Chris J. Taylor[a]

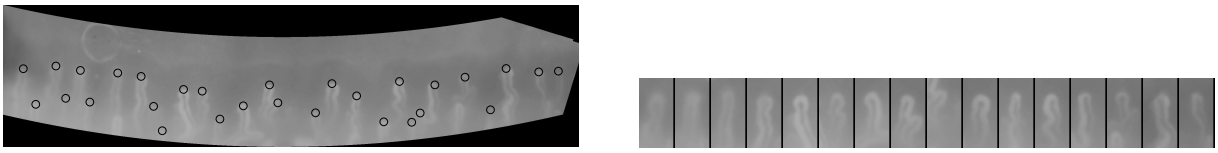
[a]Division of Imaging Science, Stopford Building, Oxford Road, University of Manchester, M13 9PT, UK,
[b]University of Manchester,Rheumatic Diseases Centre, Hope Hospital, Salford, M6 8HD, UK

Abstract. When Raynaud's phenomenon - episodic colour change of the fingers, usually in response to cold exposure - occurs secondary to the connective tissue disease systemic sclerosis (SSc), characteristic microvascular changes occur. Changes in certain features (capillary width, tortuosity, derangement and density) can be observed and measured from microscope images of nailfold capillaries. We have developed methods to enhance the images, to measure the above features, and to classify them into three categories: normal control, primary (idiopathic) Raynaud's phenomenon, and SSc. We show that the results can help to analyze the predictability of each feature change and the categorization of the Raynaud's phenomenon into primary or secondary.

1 Introduction

Nailfold capillaroscopy, which involves taking still or video photomicrography of the capillaries in the nailbed, has long been recognized as a convenient way to investigate Raynaud's phenomenon. When this occurs in patients with an underlying connective tissue disease such as systemic sclerosis (SSc), it can be associated with structural microvascular change. However, it is believed that in patients with primary (idiopathic) Raynaud's phenomenon, structural microvascular change does not occur and the phenomenon is purely vasospastic [1]. It is important to differentiate between patients with PRP and SSc because only in patients with SSc or other forms of secondary Raynaud's does the condition progress to irreversible tissue ischaemia. Almost all patients with SSc experience severe Raynaud's phenomenon, and this is the presenting feature in approximately 70% [2] of the cases. Our aim was to measure the severity of any microvascular change from capillaroscopic images, and categorise and analyse images in different subgroups of patients with Raynaud's phenomenon using quantitative measurements.

This study is based on mosaic images, which are a combination of several microscopic images showing a panoramic image of the capillaries in the nailfold, with a manual marking on the apex of each capillary [3, 4]. Based around the manual markings, we extract a 'region of interest' (ROI) for both the whole image and each individual capillary in order to focus only on capillary loops. The ROI for the whole image lies between two parallel polynomial curves above and below a "line of best fit" among the manual markings. The original images have dark capillaries on a lighter background, and we invert the images for greater ease of interpretation. The rest of the work is based on these converted ROIs, examples of which are shown in figure 1.



a) b)

Figure 1. ROI for a) Mosaic image; b) Individual Capillaries.

Changes in features, – width, tortuosity, derangement and density – are used as indicators for diagnosis. Some researchers measure the capillary dimensions (widths and lengths) and capillary density quantitatively by hand, though this is quite tedious and the measurements are subjective because different people will choose different capillary sites for measurements. Thus, there is a high demand for automatic measurements to assist in clinical analysis. We try to use artificial classifiers to classify the subjects into normal, PRP and SSc based on these measurements.

In the study used in this paper, 119 capillaroscopy images were available for image analysis, of which 48 were control subjects, 20 were PRP subjects and 51 were SSc subjects. The allocation of each subject to a disease category is based on various clinical diagnostic methods including observing their capillaroscopy images. This study was carried out in three stages: image enhancement, feature extraction and classification, which are described below.

*Email:kaiyan.feng@stud.man.ac.uk

2 Methods

In a previous paper we described a system to enhance the visibility of the capillaries and to estimate an orientation at each pixel [5]. Basically, we use rotating directional openings (openings using a linear structuring element over a range of angles) to enhance linear structures while removing the variation of the background, and the angle at which the linear opening generating the greatest value occurs is estimated as the orientation of that pixel; then we use an orientation filter to enhance those pixels whose orientations are closer to vertical, because the pixels on a capillary tend to have a vertical direction; finally, we estimate the probability of each pixel falling on a linear structure and use that probability to further enhance the images [5].

2.1 Feature Measurement

Width, tortuosity, derangement and density are measured using the final enhanced images f_{final}. Though the giant capillaries are poorly enhanced in f_{final} because they don't usually appear to be linear, the advantages still outweigh the disadvantages. This was partly demonstrated in our previous paper [5].

In [5], we showed how to measure the capillary width and tortuosity from ROIs for mosaic images ROI_m. Width is measured using pseudo-granulometry adopted from granulometry. Granulometry, developed by Matheron [6], is commonly applied to measure the size of small structures by employing solid structuring elements such as disks. The pseudo-granulometry used in this study uses annuli instead as they appear to be more robust to noise. Tortuosity is measured as the entropy of the intensity-weighted orientation histogram, which is defined as $S(ROI, \theta) = \sum_{d(x)=\theta} ROI(x)$ $(0 \leq \theta \leq \pi)$, where $d(x)$ is the orientation of a pixel x in the ROI. Similarly, we can also measure the width and tortuosity for each capillary based on ROIs for individual capillaries ROI_c. The contrast of an individual capillary Γ_c is measured as the difference between the maximum average intensity and the minimum average intensity in the series of openings in the pseudo-granulometry. The average width and tortuosity then can be measured as $\overline{W} = \sum_{i=1}^{n}(\Gamma_{c(i)} \cdot W_{c(i)}) / \sum_{i=1}^{n} \Gamma_{c(i)}$ and $\overline{T} = \sum_{i=1}^{n}(\Gamma_{c(i)} \cdot T_{c(i)}) / \sum_{i=1}^{n} \Gamma_{c(i)}$ respectively. We also measured the STD of the widths σ_W weighted by Γ_c.

Derangement indicates how much the axes of capillaries vary from each other. We first calculate the orientation of each capillary, then the derangement is calculated as the standard deviation of these orientations weighted by their contrasts. The orientation of each capillary can be calculated from its intensity-weighted orientation histogram $S(ROI_c, \theta)$ which is first decomposed into x and y coordinates as $sum(S_x(ROI_c)) = \sum_{\theta=1}^{\pi} S(ROI_c, \theta) \times \cos(\theta)$ and $sum(S_y(ROI_c)) = \sum_{\theta=1}^{\pi} S(ROI_c, \theta) \times \sin(\theta)$. Then the dominant direction of capillary c can be measured as $\theta_c = \{\theta'_c$ if $\theta'_c \geq 0$ otherwise $\theta'_c + \pi\}$, where $\theta'_c = \arctan[sum(S_y(ROI_c))/sum(S_x(ROI_c))]$, such that $0 \leq \theta_c \leq \pi/2$. The mean orientation of capillaries $c(i)(i = 1, ..., n)$ is measured as $\overline{\theta_c} = \sum_{i=1}^{n}(\Gamma_{c(i)} \cdot \theta_{c(i)}) / \sum_{i=1}^{n} \Gamma_{c(i)}$. So the derangement of the capillaries can be calculated as the standard deviation of the capillaries' dominant directions weighted by their contrasts as $D = \sqrt{(1/\sum_{i=1}^{n} \Gamma_{c(i)}) \cdot \sum_{i=1}^{n} [\Gamma_{c(i)} \cdot (\theta_{c(i)} - \overline{\theta_c})]^2}$.

The average distance between capillaries can be taken as the capillary density. We first extracted the skeletons of the capillaries and worked out the density from the distances between those skeleton points. Non-maximal suppression (NMS), originally developed by Rosenfeld [7], was used to extract the backbone of capillaries. With the NMS algorithm, each pixel is compared with its two immediate neighbors lying perpendicular to its direction, and it is selected if its magnitude is greater than or equal to that of its two neighbors. f_{final} was first smoothed by a Gaussian lowpass filter to remove high frequency noise and to enhance the capillary ridges, then the NMS algorithm was applied to the filtered image, and finally the image was processed using a thinning algorithm to extract the capillaries' skeleton. The distance between any two immediately neighboring skeleton points lying along the second polynomial curves, which run parallel to the curve defining the ROI, is called run-length l. Because two neighboring points can be taken from the same capillary, giving a shorter run-length than for those taken from different capillaries, longer run-lengths should be given more weight for density measurement. Therefore, if the histogram of all run-lengths between any two neighboring skeleton points is denoted as $h(l)$, density is calculated as $N = \sum_{l=1}^{n} h(l) \cdot d^2$ which is the second moment of $h(l)$.

2.2 Classification

We have chosen linear SVMs (support vector machine) [8] as the basic classifiers because they have good generalization (the ability to classify new data correctly). But an SVM classifier is a binary classifier (it can only separate two classes), a set of SVM classifiers need to be combined together to solve a multi-class problem. Two common strategies to generate pairs of classes are one-vs-others classes: control vs PRP+SSc($CvsO$), PRP vs control+SSc($PvsO$), and SSc vs control+PRP($SvsO$); and all-pairs classes: control vs PRP($CvsP$), control vs SSc($CvsS$), and PRP vs SSc($PvsS$). The pair of classes $PvsO$ are excluded because PRP data lie between control and SSc data which can not be separated by a linear SVM, leaving only 5 pairs of classes available to be used.

For each pair of classes, SVMs were established by selecting subsets of $\{\overline{W}, \sigma_W, \overline{\theta_c}, \overline{T}, D, N\}$ as its input features. Each SVM was evaluated by 4-fold cross-validation, i.e. the data were divided into 4 folds randomly 10 times, each time every 3 folds of data were trained by SVM and the remaining fold of data were tested to evaluate the SVM performance. Because an SVM is a binary classifier, the test data were evaluated using an ROC (Receiver operating characteristic) curve [9]. Each hyperplane h parallel to the separating hyperplane obtained by training the SVM will generate a true positive TP_h and a false positive FP_h from the test data, corresponding to a point (FP_h, TP_h) on the ROC curve. The area under an ROC curve A_z was used to assess the accuracy of the classifier. So for each feature subset, 10×4 ROC curves and their A_z, the mean $\overline{A_z}$ and the standard error σ_{A_z} were obtained by training and evaluating the SVMs. The feature subset which generated the highest value $\overline{A_z}$ was used to separate the associated pair of classes. Finally, given a training data set, SVMs used the optimized feature subsets to train the five pairs of classes, generating a separating hyperplane. In order to achieve minimum classification error, we chose a hyperplane h (parallel to the separating hyperplane given by SVM), which would let $TP_h - FP_h = arcmax(TP - FP)$, as the hyperplane to separate new data. We denote these classifiers as $f_1, f_2, ..., f_5$ for the five pairs of classes $CvsO$, $SvsO$, $CvsP$, $CvsS$ and $PvsS$ respectively.

Classifiers $f_1, f_2, ..., f_5$ were combined using a loss-based decoding method [10] to build a multi-class classifier F. First a prediction coding matrix $M \in \{-1, 1\}^{3\times5}$, which is used to specify the relationship between the SVMs and the categories, was constructed with the three categories as its rows and the five SVMs as its columns. An element in M was allocated a value of 1 (-1) if the category in the row was used as a positive (negative) category to train the SVM in the column and 0 if the SVM was not used for the category's classification. A given data point $\overline{x} = \{\overline{W}, \overline{T}, D, N\}$ was classified using $f_1, f_2, ..., f_5$ respectively, generating a classification result $f(\overline{x}) = \{f_1(\overline{x}), f_2(\overline{x}), ..., f_5(\overline{x})\}$ which was used to calculate its distance with each row vector in M. The distance between r^{th} row vector $M(r)$ and $f(\overline{x})$ was calculated as $dis(M(r), f(\overline{x})) = \sum_{b=1}^{5} L(M(r,b) \cdot f_b(\overline{x}) \cdot A_z(f_b))$, where $L(z) = \{0 \text{ if } z > 1 \text{ otherwise } (1-z)\}$ is a loss function. Data point $\overline{x}$ belongs to r^{th} class if the associated distance is shortest among the three classes. The performance of the multi-class classifier was evaluated by 4-fold cross-validation. The use of cross-validation was similar to that in evaluating the SVMs, except that instead of an ROC curve the test data was used to obtain a correct classification rate C_r and its standard error.

3 Results and Discussion

Let $A_z(pc, fs)$ and $C_r(pc, fs)$ be the A_z and C_r obtained by evaluating classifier SVM using 4-folds cross-validation to classify a pair of classes pc with a feature subset fs. A_z and C_r obtained by using best single feature or best combined features for SVMs are shown in table 1. The confusion matrix of the results of the multi-class classifier F is shown in table 2. Besides automatic measurements, we also used manual measurements: manual mean width $\overline{MW}$ (described in paper [5]) and manual density $\overline{MN}$(the average distance between the capillary apex markings) for classification. We will compare some of the results to those of the automatic method.

Table 1 shows that for a single feature: 1. mean width and density are the most useful features for classification, 2. mean width is better than width STD for classification, which may indicate that all capillaries in the nailfold are affected during vasospastic attacks, 3. mean width and tortuosity are better in separating control and PRP, which may mean that they are the first signs of Raynaud's phenomenon, 4. density is the best feature to separate SSc from either PRP or control, which may indicate that loss of capillaries is the worst symptom for Raynaud's phenomenon in capilloroscopy, 5. automatic width is more effective in separating control and PRP than the manual width ($A_z(CvsP, \overline{W}) = 0.7442 > A_z(CvsP, \overline{MW}) = 0.6925$), 6. automatic density is more effective in separating control and PRP than manual density ($0.7204 > 0,5971$), but less effective in separating PRP and SSc ($0.7885 < 0.9790$), which may well be because they measure density differently; for combined features: 1.

Notations	$CvsO, SvsO, CvsP, CvsS, PvsS$:$G1, G2, G3, G4, G5$; $\overline{W}$:1, σ_W:2, $\overline{\theta_c}$:3, $\overline{T}$:4, D:5, N:6				
Best Single Feature for SVM to Obtain A_z and C_r					
pc, fs	$G1$, {6}	$G2$, {6}	$G3$, {1}	$G4$, {6}	$G5$, {6}
A_z	0.8377 ± 0.0101	0.8559 ± 0.0101	0.7442 ± 0.0188	0.8839 ± 0.0097	0.7885 ± 0.0159
pc, fs	$G1$, {1}	$G2$, {6}	$G3$, {4}	$G4$, {6}	$G5$, {6}
$C_r(\%)$	73.19 ± 1.04	78.82 ± 0.84	76.47 ± 1.30	80.00 ± 1.08	72.12 ± 1.39
Best Combined Features for SVM to Obtain A_z and C_r					
pc, fs	$G1$, {1, 3, 6}	$G2$, {1, 5, 6}	$G3$, {1, 3, 6}	$G4$, {1, 6}	$G5$, {2, 4, 5, 6}
A_z	0.9142 ± 0.0058	0.9232 ± 0.0051	0.7704 ± 0.0197	0.9549 ± 0.0055	0.8494 ± 0.0084
pc, fs	$G1$, {1, 2, 4, 6}	$G2$, {1, 5, 6}	$G3$, {1, 5}	$G4$, {2, 4, 6}	$G5$, {1, 5, 6}
$C_r(\%)$	84.29 ± 0.88	86.55 ± 0.84	79.26 ± 1.41	88.18 ± 1.05	77.32 ± 1.50

Table 1. The Best Classification Results Obtained from SVMs.

	Control (%)	PRP (%)	SSc (%)
Control	72.080 ± 2.360	23.330 ± 2.176	4.583 ± 0.941
PRP	31.000 ± 3.199	46.500 ± 2.809	22.500 ± 2.395
SSc	3.910 ± 0.738	14.540 ± 1.630	81.550 ± 1.785

Table 2. The Confusion Matrix of the Multi-class Classifier F.

Besides width and density which are apparently the most useful features for combination, derangement is quite a good complement, 2. width and tortuosity don't usually appear together in a combination, which may indicate that they change together as the disease progresses, 3. manual measurements alone are very good at separating SSc from the others, but they are less effective than the automatic measurements in separating control from PRP, 4. though some features correlate significantly with each other, but the feature combination still improves the classification.

Table 2 shows that 1. control and SSc can be separated from each other quite well, which confirms that SSc is associated with structural microvascular change, 2. PRP is very hard to separate from control, which indicates that many of the PRP subjects are benign, 3. PRP lies between control and SSc, but tends to lean more towards the control group, 4. 'control' may not be an ideal control group, we suggest that besides vasospastic attacks, some other criteria such as the extremities' resistance to cold should be used to define an ideal control group.

In summary, we have described a method for automatically classifying patients into normal and two separate disease groups on the basis of their nailfold images. The results are clinically useful, given the difficulty of the task for clinicians, and a range of classification results can be obtained using different combinations of the binary classifiers.

References

1. E. C. LeRoy & T. A. Medsger. "Raynaud's phenomenon: a proposal for classification." *Clin Exp Rheumatol* **10**, pp. 485–288, 1992.
2. J. J. P. Belch. "Raynaud's phenomenon: its relevance to scleroderma." *Annals of the Rheumatic Diseases* **50**, pp. 839–845, 1991.
3. P. D. Allen, C. J. Taylor, A. L. Herrick et al. "Enhancement of temporally variable features in nailfold capillary patterns." In 9^{th} *British Machine Vision Conference*, pp. 698–705. 1998.
4. P. D. Allen, C. J. Taylor, A. L. Herrick et al. "Image analysis of nailfold capillary patterns from video sequences." In 2^{nd} *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 698–705. 1999.
5. K. Feng, P. D. Allen, T. Moore et al. "Automatic capillary measurement." In *MIUA'2003*, pp. 161–164. 2003.
6. G. Matheron. *Elements pour une Theorie des Milieux Poreux (in French)*. Masson, Paris, 1967.
7. A. Rosenfeld & A. C. Kak. *Digital Picture Processing*. Academic Press, New York, 1982.
8. S. R. Gunn. "Support vector machines for classification and regression." Technical report, Image Speech and Intelligent Systems Research Group, University of Southampton, 1997.
9. M. H. Zweig & G. Campbell. "Receiver-operating characteristic (roc) plots: a fundamental evaluation tool in clinical medicine." *Clinical Chemistry* **39**, pp. 561–577, 1993.
10. E. L. Allwein, R. E. Schapire & Y. Singer. "Reducing multiclass to binary: A unifying approach for margin classifiers." *Journal of Machine Learning Research* **1**, pp. 113–141, 2000.

Detailed Comparison of Non-Rigid Registration in Inter-subject Brain Registration

William R. Crum[a], Daniel Rueckert[b], Mark Jenkinson[c], David Kennedy[d] and Stephen M. Smith[c*]

[a]Division of Imaging Sciences, Kings College London, [b]Dept. of Computing, Imperial College, London, [c]FMRIB, Oxford, [d]Center for Morphometric Analysis, Department of Neurology, MGH, USA

Abstract A growing number of neuroimaging studies rely on increasingly sophisticated non-rigid registration schemes. Validation of non-rigid registration, particularly with relation to specific neuroanatomy remains an unsolved problem. In this paper, densely labelled magnetic resonance brain images from 8 subjects are used to assess three non-rigid registration algorithms (affine (FLIRT), free-form deformation (B-Spline) and fluid). Ninteen meta-labels are constructed from the original label set chosen to sample a range of scales and locales in the brain. Each subject is registered to all other subjects to prevent reference bias. Each registration algorithm is assessed for its ability to bring the labels into correspondence over all subjects. A simple label overlap measure is used to assess the correspondence. Results show that for this data and these algorithms, registration using free-form deformations performed best on average. All algorithms varied significantly in their ability to align different structures however which may have implications for contemporary neuroimaging studies. Overlaps of label pairs were correlated to varying degrees which should be accounted for in future evaluation studies.

1 Introduction

Validation and performance assessment of non-rigid registration in medical applications remains difficult because of the lack of a ground truth in many cases. Often non-rigid registration is applied to achieve a specific goal such as to bring homologous structures into alignment; the criterion for success is whether the edges of such structures are aligned as often there is little internal structure visible in imaging on which to base an assessment. The most common example of this is in inter-subject brain registration for atlas construction and population-based morphometry studies. In these cases an appropriate validation is one based on structural labels that allow the degree of correspondence pre- and post-registration to be determined using label overlap measures. In this paper we apply an evaluation framework of this kind to three registration algorithms that have all been used in brain registration and identify structures that fail to be well registered by any algorithm. It is important to understand that this evaluation does not test whether any algorithm achieves good point-to-point correspondence but tests whether identifiable anatomical features are brought into the same volume of space by registration; (indeed for inter-subject brain registration it is doubtful that a point-to-point correspondence exists and unlikely that it could be established from MR registration alone if it did exist [1]). This evaluation allows algorithms to be compared purely on the basis of their ability to accomplish the task. It also allows an evaluation on the basis of which brain structures or tissue classes are consistently well registered or poorly registered; this may have implications for large neuroimaging studies where inferences about structure and morphology within and across populations are made on the basis of non-rigid registration results.

2 Methods

2.1 Brain Labels

Eight labelled MR brain images were obtained from the Centre for Morphometric Analysis at MGH (Boston). Each voxel has associated binary labels that denote a particular structure or tissue class; there are 84 sub-cortical labels and 48 cortical labels. These brain images and the sub-cortical labels are available from the Internet Brain Segmentation Repository (http://www.cma.mgh.harvard.edu/ibsr/) for use by the research community; image numbers 1-5, 7, 9, and 10 of the cohort were used in this work. The original labels were selectively grouped to produce a smaller set ranging in scale from the entire brain and the primary lobes down to structures such as the

*Corresponding Author: steve@fmrib.ox.ac.uk, Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, UK

hippocampus and the thalamus (see Table 1 for the full list of labels). The chosen groupings are arbitrary and can be defined to suit any specific application. We refer to the set of grouped labels as the test labels.

Major Structures	Major Lobes	Other Structures	Sub-Cortical
All (32)	Frontal (14)	Lat. Ventricle (2)	Thalamus (2)
Brain (25)	Occipital (8)	Cerebellum (4)	Caudate (2)
Cortex (6)	Parietal (7)	Brain Stem (3)	Putamen (2)
White Matter (2)	Temporal (16)	Sub-Cortex (10)	Pallidum (2)
CSF (7)			Hippocampus (2)
			Amygdala (2)

Table 1. The four groups of test labels used in the evaluation. The number of original anatomical labels used to create each test label is shown in brackets. All structures except for the Major Lobes are constructed with data publically available for research purposes from the Internet Brain Segmentation Respository.

2.2 Registration Algorithms

Three non-rigid registration algorithms were evaluated: (a) an affine registration algorithm (FLIRT, part of the FSL image analysis toolkit (www.fmrib.ox.ac.uk/fsl)) [2], [3] (b) a free-form deformation algorithm based on B-Splines [4], [5] and (c) a viscous fluid algorithm [6] implemented following [7].

FLIRT: (FMRIB's Linear Registration Tool) FLIRT is highly robust to the initial image alignment by use of a customised global optimisation method that runs over multiple scales (8, 4, 2 and 1mm); a large search space in rotation and scale is used at the 8mm scale and many smaller perturbations on the best three candidate solutions applied at 4mm resolution. Contributions from near the edge of the overlapping field of view are de-weighted in order to produce a smoothly changing cost function. The correlation ratio is used to drive the registration.

B-SPLINE: Non-rigid registration based on free-form deformations (FFDs) and B-Splines is widely used in many registration applications including those involving inter-subject brain registration. The basic idea of FFDs is to deform an object by manipulating an underlying mesh of control points. The optimal control point locations are found by minimising a two-part cost function; one term is a voxel-based similarity measure, (in this case normalised mutual information), and the other term constrains the transformation to be smooth. The control-point spacing was set at 2.5mm for this study.

FLUID: The fluid methods use a mathematical model of a compressible viscous fluid for the transformation between images. They can be less robust than other methods without good initialization but can accommodate large deformations without folding of the displacement field. Here the fluid algorithm was initialised from a locally derived affine registration of each subject into a standard anatomical space. Unlike the other two algorithms the fluid registration was run at single resolution. A multi-resolution fluid algorithm will be evaluated in the future. The intensity cross correlation was used to drive the registration.

2.3 Overlap Measures

There is considerable literature on the subject of assessing overlap between two labels, S and T, much of it applied to the assessment of segmentation algorithms (e.g. [8]). In this work we use a conservative measure, P, the ratio of the number, N(), of overlapping voxels, to the total number of voxels in the labelled structures.

$$P = \frac{N(S \cap T)}{N(S \cup T)} \quad (1)$$

2.4 Evaluation Framework

Each of the eight test subjects was registered to the other seven subjects using each of the three algorithms giving 56 inter-subject registrations for each algorithm. The registrations were run independently by the researchers most familiar with their operation using their standard parameter choices. Each test label on each subject was transformed onto every other subject using the results of each registration algorithm and trilinear interpolation thresholded at 50% to produce transformed binary labels. The test-label overlap, P, was computed in all cases and this data was analysed to produce a mean and standard deviation fractional overlap for each test-label for each registration algorithm. The overlaps were also computed for all pairs of unregistered scans.

3 Results

The FLIRT registrations took approximately 5 minutes each to run on a contemporary desktop Linux PC. Both the B-Spline and fluid registrations were run in a distributed fashion on a Linux condor cluster (one CPU per registration) and typically took between 2 and 10 hours per registration. The results are summarised in Fig. 1. The B-Spline method performed best overall. The "all" "brain" and "cerebellum" labels scored the highest mean overlaps over all algorithms. The largest range in mean overlaps between techniques could be seen in the "csf" [0.35, 0.73], "lv" [0.37, 0.78] and "cortex" [0.48, 0.73] labels. Some structures remained poorly registered by all techniques e.g. "amydala" [0.35, 0.47] and "hippocampus" [0.41, 0.53]. In all cases all registration algorithms improved the mean label overlap except for FLIRT applied to the pallidum and amygdala. We might expect that the potential accuracy of matching increases with the degrees of freedom available to the algorithms but that the potential for mis-registration also increases. What we found in this study is that the B-Spline method consistently performed well and was only outperformed by the fluid method for two structures (the pallidum and amygdala) where all methods struggled to improve the overlap. Conversely the B-Spline method proved superior to the other methods for registering the larger tissue compartments (cortex, white, csf and lateral ventricle). There are some caveats associated with these results. The most obvious is that the sample-size of subject data is relatively small but another important factor is that the images were rotationally realigned as part of the labelling process, and had already been interpolated prior to our analysis.

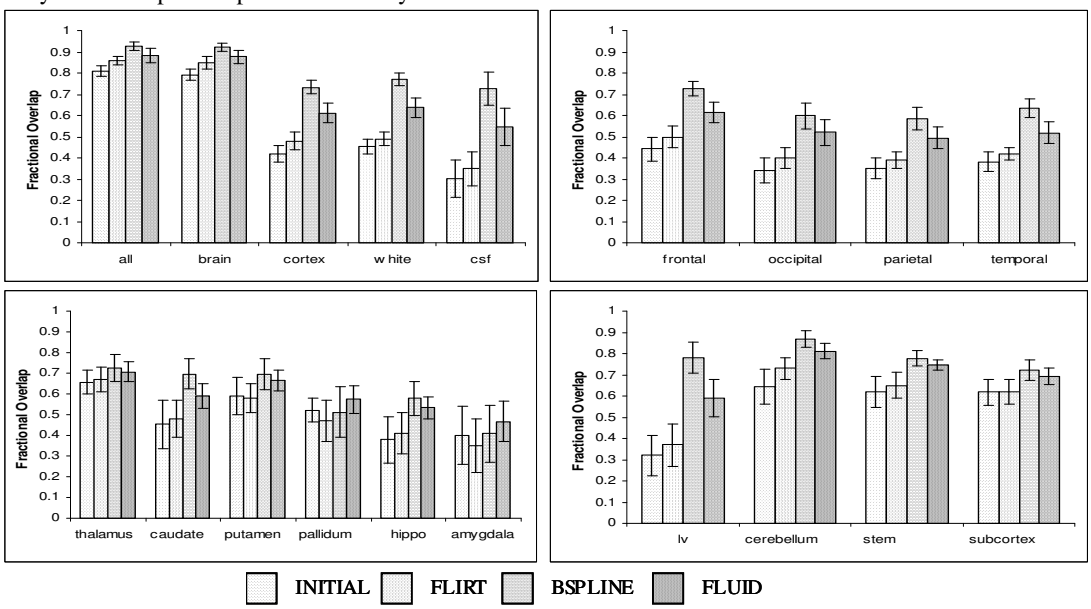


Fig. 1. The mean and standard deviation of the overlap measure for each structure registered by each registration algorithm. INITIAL refers to the original images, which had been rotationally realigned as part of the labelling process.

4 Discussion

We have assessed the ability of three registration algorithms to align a variety of brain structures between eight subjects and found overlaps ranging from ~0.3 (lateral ventricle) to ~0.9 (brain). The results enable us to distinguish the performance of the algorithms over different parts of the brain. Notable is that the B-Spline algorithm matches CSF and lateral ventricle particularly well, no algorithms match the major lobes better than ~0.7 and that there is less difference between the algorithms ability to match the sub-cortical structures than a consideration of the degrees of freedom of the transformation model used in each case might suggest. We would normally expect the largest increase in overlap to be achieved by the affine registration compared with no registration however due to the realignment of the images prior to labelling the observed increases are relatively small. The fact that each algorithm used a different image similarity measure may impact on the results but also reflects the current lack of consensus and deep understanding of the operation of such measures.

The most relevant recent work on registration evaluation is [9] where six registration algorithms were compared using a variety of measures including tissue overlap, correlation of differential characteristics and sulcal shape characteristics. They found that algorithms with higher degrees of freedom did not perform

proportionately better at matching cortical sulci and that inter-subject cortical variability remains a severe challenge for voxel-based non-rigid algorithms. These findings are consistent with our experience. Previously, Grachev *et al* [10] suggested using 128 carefully defined and manually placed landmarks per hemisphere to evaluate inter-subject registration accuracy. Maudgill *et al* [11] identified 24 homologous cortical surface landmarks that could either be used for registration assessment or to drive registration. Landmarks can provide a mm error estimate but are not easily related to anatomically important structures. In the future, hybrid assessments based on landmark and regional assessment may prove useful.

The overlap measure we use in this paper is well known but does not account explicitly for inconsistency in the labelling process nor does it indicate the nature of the error in non-perfect overlaps. Crum *et al* [1] suggest the use of a tolerance parameter with overlap measures so that, for instance with the tolerance set to 1 voxel, the boundaries of two labels are regarded as completely overlapping if they are at most one voxel away from each other. A deeper consideration of overlap measures and the incorporation of labelling error is an urgent priority. When evaluating overall performance with a set of labels there may be correlation between labels most obviously because they share a common boundary (e.g. amygdala and hippocampus) or because one is a subset of the other (e.g. grey-matter and brain). We computed correlation of overlaps between pairs of labels and found that the correlations exist and are not ranked the same for each algorithm. For instance the correlations between lateral ventricle overlaps and caudate overlaps for FLIRT, B-Spline and fluid were 0.39, 0.70 and –0.13 respectively but for putamen and sub-cortex were 0.84, 0.71 and 0.69 respectively. This reflects interactions between the transformation models employed by each algorithm and the scales of image structure driving the registration.

We plan to extend this study by including more registration algorithms and more labelled subjects in the evaluation and more interestingly to build a more robust framework for assessing performance using collections of labels. Future work will focus on a more detailed technical comparison of algorithms but this initial work has provided a benchmark for future performance in two ways. First, other registration techniques can be easily tested within the same framework for an operational comparison. Second, we can recognise that these algorithms are subject to many parameter choices that we have ignored in this study. We can optimise the parameter choice for each algorithm with respect to this well-defined task; this optimisation process may ultimately lead to new methods tuned to register structures of particular scale and intensity characteristics.

Acknowledgements

The authors are grateful for the intellectual and financial support of the Medical Images and Signals IRC (EPSRC and MRC GR/N14248/01). This work was made possible by the research environment created by this consortium. We acknowledge valuable contributions from Professor David Hawkes and Dr Derek Hill.

References

1. W.R. Crum, L.D. Griffin, D.L.G. Hill, and D.J. Hawkes "Zen and the Art of Medical Image Registration: Correspondence Homology and Quality", *NeuroImage* **20**, pp 1425-1437, 2003.
2. M. Jenkinson & S.M. Smith "A Global Optimisation Method for Robust Affine Registration of Brain Images", *Medical Image Analysis* **5(2)**, pp 143-156, 2001.
3. M. Jenkinson, P. Bannister, J.M. Brady, and S.M. Smith "Improved Optimisation for the Robust and Accurate Linear Registration and Motion Correction of Brain Images", *NeuroImage* **17(2),** pp 825-841, 2002.
4. D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach and D.J. Hawkes "Non-Rigid Registration Using Free-Form Deformations: Application to Breast MR images", *IEEE Transactions on Medical Imaging* **18(8)**, pp 712-721, 1999.
5. D. Rueckert, A.F. Frangi, and J.A. Schnabel, "Automatic Construction of 3D Statistical Deformation Models of the Brain using Non-Rigid Registration", *IEEE Transactions on Medical Imaging* **22(8)**, pp 1014-1025, 2003.
6. G.E. Christensen, S.C. Joshi and M.I. Miller, "Volumetric Transformation of Brain Anatomy", *IEEE Transactions on Medical Imaging* **16(6)**, pp 864-877, 1997.
7. W.R. Crum, R.I. Scahill, N.C. Fox, "Automated Hippocampal Segmentation by Regional Fluid Registration of Serial MRI : Validation and Application in Alzheimer's Disease", *NeuroImage* **13(5)**, pp 847-855, 2001.
8. G. Gerig, M. Jomier and M. Chakos, "Valmet: A new validation tool for assessing and improving 3D object segmentation" In *Medical Image Computing and Computer-Assisted Intervention*. Lecture Notes in Computer Science **2208**, Springer-Verlag, Berlin Heidelberg New York, pp 516-528, 2001.
9. P. Hellier, I. Barillot, B. Corouge, G. Gibaud, G. Le Goualher, D.L. Collins, A. Evans, G. Malandain, N. Ayache, G.E. Christensen and H.J. Johnson, "Retrospective Evaluation of Intersubject Brain Registration", *IEEE Transactions on Medical Imaging* **22(9)**, pp 1120-1130, 2003.
10. I.D. Grachev, D. Berdichevsky, S.L. Rauch, S. Heckers, D.N. Kennedy, V.S. Caviness and N.M. Alpert, "A Method For Assessing the Accuracy of Intersubject Registration of the Human Brain Using Anatomic Landmarks", *NeuroImage* **9**, pp 250-268, 1999.
11. D.D. Maudgil, S.L. Free, S.M. Sisodiya, L. Lemieux, F.G. Woermann, D.R. Fish, and S.D. Shorvon, "Identifying Homologous Anatomical Landmarks On Reconstructed Magnetic Resonance Images of the Human Cerebral Cortical Surface", *Journal of Anatomy* **193**, pp 559-571, 1998.

Invisible differences in brain images of schizophrenics and controls

Vassili A Kovalev and Maria Petrou

Centre for Vision, Speech and Signal Processing, School of Electronics and Physical Sciences
University of Surrey, Guildford GU2 7XH, United Kingdom
{v.kovalev,m.petrou}@surrey.ac.uk

1 Introduction

It is commonly accepted that the symptoms of schizophrenia have their origins in disordered brain chemistry, diffuse cortical abnormalities and that the condition is also associated with certain morphological brain changes (e.g. [1], [2], [3]). The methodological base of neuroimaging research of schizophrenia includes conventional volumetric measurements (e.g. [4]), comparison of the differences in relaxation times between the two groups (e.g. [5], [1]), measuring MR signal hyperintensities (e.g. [6]), voxel-based studies that help to identify structural changes in the brains of schizophrenics [3], detection of structural differences between the brains using 3D volumetric texture analysis [7], as well as some special methods (e.g. [8]).

It is possible that interesting brain tissue variations between groups of subjects may be observed if differences in higher order statistics are examined. The higher order statistics are often considered in two different contexts of image analysis. First, to denote the *spatial* statistics of image change [9], [10]. Consideration of the image spatial domain is particularly relevant to such image patterns as 2D and 3D textures (e.g. [11], [7]). We are not interested here in texture in this commonly understood sense of the word, i.e. in spatial patterns. Instead, we are interested in regions which look rather random and their only difference is *the underlying probability density function* from which the observed values are drawn. More specifically, we are interested in detecting boundaries between regions the histograms of which differ in their third order statistics only. These boundaries are of particular interest because it is known that the human vision system can only see the boundaries between regions that differ in their first or second order statistics. Thus, in this paper we identify boundaries between brain regions which differ in third (skewness) but not in the first (mean) and second (variance) order statistics and compare the extent of such regions in schizophrenic patients and normal controls.

2 Materials and Methods

The MRI-T2 and MRI-PD data used refer to 40 subjects (21 schizophrenic patients and 19 normal controls). All subjects were scanned with a 1.5-T GE Signa (GE medical systems Milwaukee) at the Maudsley Hospital, London. Proton density and MRI-T2 weighted images were acquired with a dual-echo fast spin-echo sequence (TR=4000 ms, TE=20.85 ms). Contiguous interleaved images were calculated with 3-mm slice thickness and 0.856×0.856 mm in-plane pixel size in the axial plane parallel to the intercommissural line. The two groups were matched for age and social class as defined by the occupation of the head of the household at the time of birth. (Mean age for controls 31.5, with a standard deviation of 5.9. Mean age for patients 33.7 with a standard deviation of 6.9. Social class 1-3, controls 18/21 and patients 16/19.) The mean premorbid IQs for both groups were estimated by the National Adult Reading Test with the normal controls having slightly higher IQ. (Mean IQ of schizophrenics 101.5 with a standard deviation 11.3 and mean IQ for normal controls 107.4 with a standard deviation 9.6.) The schizophrenics had an average of 13.1 years of education, while the normal controls had an average of 14.3 years of education. All patients had been taking typical antipsychotic medication for more than 10 years. All subjects were righthanded male. The dual-echo sequence used for image acquisition is that commonly used as part of neurological examinations at the Maudsley Hospital, London, UK. In all cases the images were preprocessed so that only the brain parenchyma was extracted for further analysis [12].

The method is based on the mapping of invisible boundaries using a spherical sliding window (Fig. 1). At every position the window was divided into two halves along the x, y and z axes. Inside each half of the window the first, second, and third moments of the data were computed. The values produced in the window halves were compared and the differences along the three axes were treated as components of gradient vectors of the statistical moments. If the magnitude of mean and variance gradients at a window position were zero within a certain tolerance, the skewness magnitude was added to all voxels of the map situated within the window. Thus the map accumulated the skewness gradient magnitude of voxels *that had zero mean and variance gradients*. The sum of these skewness gradient values were used as a feature characterising a brain and the null hypothesis that "this feature is the same for

schizophrenics and normal controls" was tested. The inter-group differences of the maps were assessed statistically using linear multi-variate techniques commonly accepted in neuroscience.

3 Results

Synthetic images. Results of detecting invisible boundaries of synthetic 3D sphere and cube images that differ from background by skewness parameter only are shown in Fig. 2. The mapping technique used in this simulated data study was as described in the previous section. The only exception was that gradients of the mean and variance values were zero by definition and the mapping condition given on the right side of Fig. 1 was not examined.

Real images. For the real data the skewness gradient maps were computed for both groups and both image modalities with window diameter varying from $d = 9$ to $d = 15$ voxels. No significant inter-group differences were found using MRI-PD data while in MRI-T2 scans the percentage of mapped brain volume was significantly lower in patients (47.5 versus 55.4 in the control group, $t = 4.57$, $p < 0.0001$, for $d = 11$). Examples of resultant skewness gradient maps superimposed onto original MRI-T2 images are shown in Fig. 3. Statistical significance plots for mapped brain volume are provided in Fig. 4. The differences between the schizophrenic and control groups were significant for all tested window diameters (the lowest significance was 48.9 versus 55.6, $t = 3.56$, $p = 0.001$ for sliding window diameter $d = 15$). Furthermore, the inter-group differences remain significant when the computed gradient is assigned to the central voxel of the window only, as opposed to all the voxels of the window (2.85% versus 3.73% in controls, $t = 3.34$, $p < 0.0019$, $d = 11$) as well as when the average skewness gradient is computed over all voxels with non-zero skewness gradient (but zero mean and variance gradients) (1.66 versus 1.69 in controls, $t = 3.62$, $p < 0.0009$, $d = 11$). However, when repeated for brain cerebral hemispheres separately, these measurements did not reveal any significant asymmetries.

In our earlier study [7] we reported structural differences between the brains of schizophrenic patients and controls detected on the same image data using 3D texture analysis methods. A comparison of these structural features with skewness gradient maps considered in the present study suggests that the new features introduced here are of a different nature than the other ones: structural differences [7] were localised in grey matter, close to the cortical surface while regions mapped here are predominantly in the white matter. A further investigation is necessary before firm conclusions can be drawn.

Acknowledgements. This work was supported by an EPSRC portfolio grant.

References

1. A. Pfefferbaum, E. V. Sullivan, M. Hedehus et al. "Brain gray and white matter transverse relaxation time in schizophrenia." *Psychiatry Research: Neuroimaging* **91**, pp. 93–100, 1999.
2. Y. Okazaki. "Morphological brain imaging studies on major psychoses." *Psychiatry and Clinical Neurosciences* **52**, pp. S215–S218, 1998.
3. T. Sigmundsson, J. Suckling, M. Maier et al. "Structural abnormalities in frontal, temporal and limbic regions and interconnecting white matter tracts in schizophrenic patients with prominent negative symptoms." *American Journal of Psychiatry* **158**, pp. 234–243, 2001.
4. S. M. Lawrie & S. S. Abukmeil. "Brain abnormality in schizophrenia. a systematic and quantitative review of volumetric magnetic resonance imaging studies." *British Journal of Psychiatry* **160**, pp. 179–186, 1998.
5. T. Supprian, E. Hofmann, M. Warmuth-Metz et al. "MRI T2 relaxation times of brain regions in schizophrenic patients and control subjects." *Psychiatry Research: Neuroimaging* **75**, pp. 173–182, 1997.
6. P. Sachdev & H. Brodaty. "Quantitative study of signal hyperintensities on T2-weighted magnetic resonance imaging in late onset schizophrenia." *American Journal of Psychiatry* **156**, pp. 1958–1967, 1999.
7. V. A. Kovalev, M. Petrou & J. Suckling. "Detection of structural differences between the brains of schizophrenic patients and controls." *Psychiatry Research: Neuroimaging* **124**, pp. 177–189, 2003.
8. C. Gaser, H. P. Volz, S. Kiebel et al. "Detecting structural changes in whole brain based on non-linear deformation application to schizophrenia research." *NeuroImage* **10**, pp. 107–113, 1999.
9. M. G. A. Thomson & D. Foster. "Role of second- and third-order statistics in the discriminability of natural images." *J. Optical Society of America* **14**, pp. 2081–2090, 1997.
10. D. Ellemberg, T. L. Lewis, K. S. Meghji et al. "Comparison of sensitivity to first- and second-order local motion in 5-years-old and adults." *Spatial Vision* **16**, pp. 419–428, 2003.
11. J. Yellott. "Implication of triple correlation uniqueness for texture statistics and the julesz conjecture." *J. Optical Society of America* **A10**, pp. 777–793, 1993.
12. J. Suckling, M. J. Brammer, A. Lingford-Hughes et al. "Removal of extracerebral tissues in dual-echo magnetic resonance images via linear scale-space features." *Magnetic Resonance Imaging* **17**, pp. 247–256, 1999.

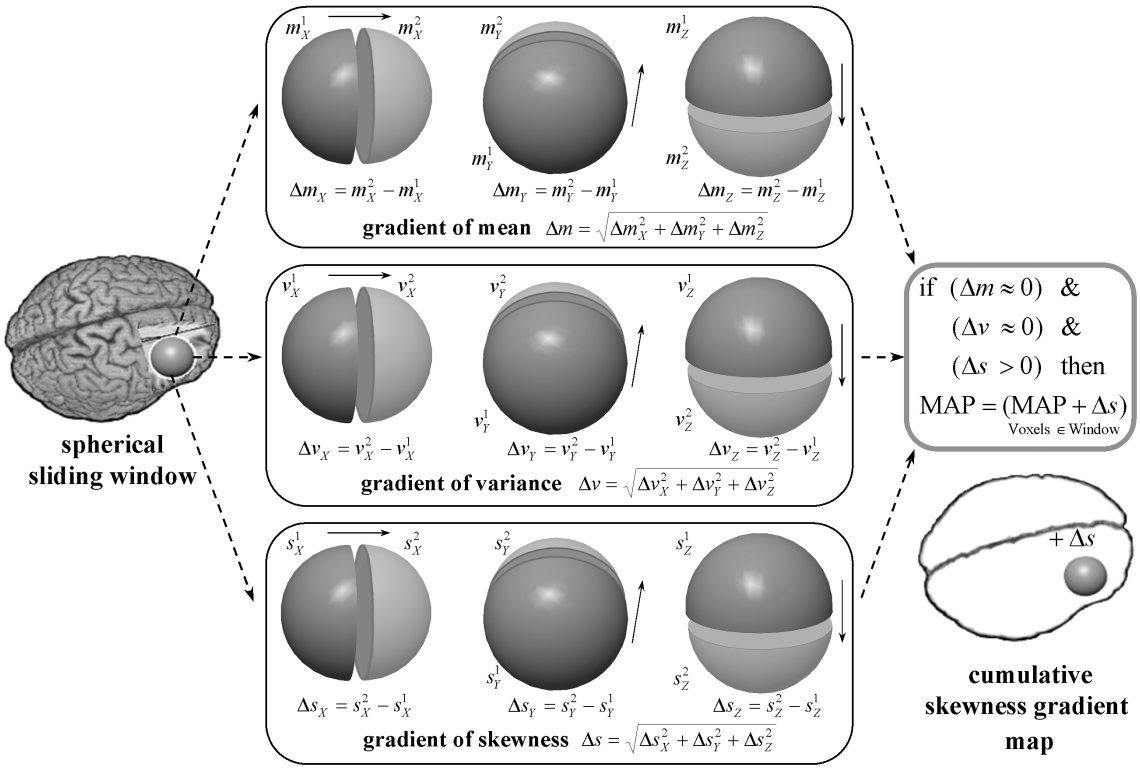


Figure 1. Mapping boundaries between brain regions which differ in the third (skewness) but not in the first (mean) and second (variance) order statistics.

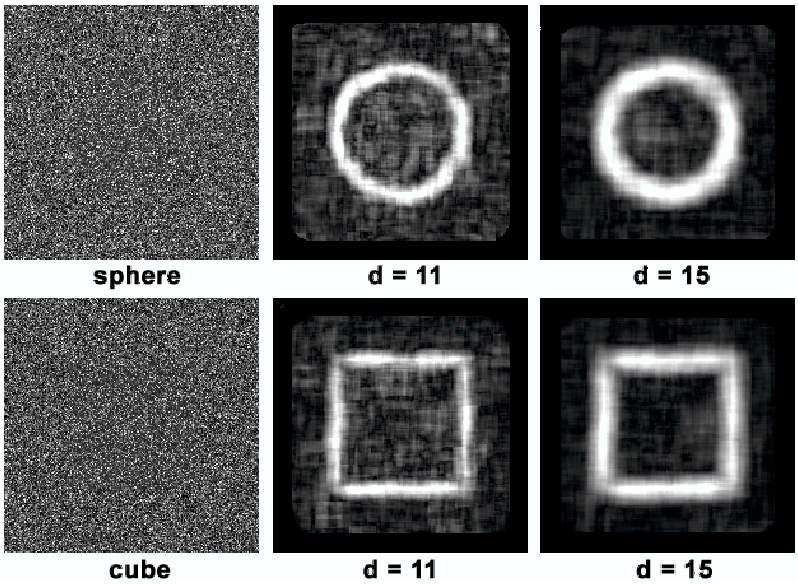


Figure 2. Mapping skewness gradient magnitude of synthetic sphere (top row) and cube (bottom row) images with different diameters of sliding window (d=11 and d=15). Image size: 200×200×200 voxels. Sphere and cube size: 101 voxes, centered. Intensity parameters: Pearson-III distribution with mean m=100, variance v=400, skewness s=1.0 for background and s=2.0 for sphere and cube. For each image the axial slice 101 is shown.

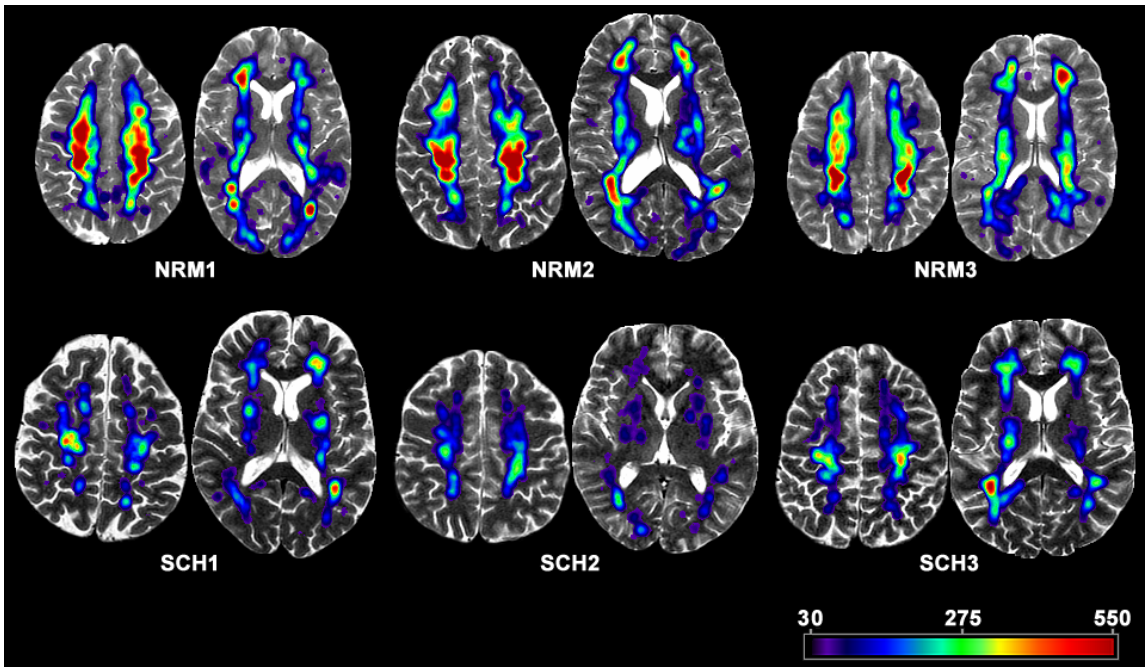


Figure 3. Examples of colour-coded skewness gradient magnitude maps for three schizophrenic patients (top row) and three controls (bottom row) superimposed onto original MRI-T2 images.

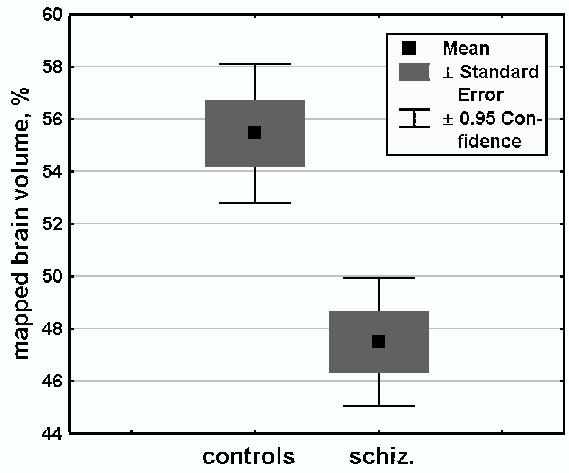


Figure 4. Statistical plots illustrating significant differences in mapped brain volume between the groups of schizophrenic patients and controls.

Simulation of Realistic Abnormal SPECT Brain Perfusion Images; Method and Applications

Tony Ward[a1], John S. Fleming[a], Sandra M. A. Hoffmann[a] and Paul M. Kemp[b]

[a]Department of Medical Physics and Bioengineering, [b]Department of Nuclear Medicine

Southampton University Hospital Trust, Southampton, UK

A method is presented that allows the fast simulation of abnormal distributions of cerebral perfusion through the synthesis of realistic stereotactic hypoperfusion lesions inside of a known normal brain. Abnormalities were first defined in standard space, as voxels with percentage perfusion reduction in known anatomical positions. The abnormality was spatially transformed to the subject space, and convolved with a measured point spread function (PSF), and then introduced into the normal image. The method is applied to a group of normal subjects to generate two sets of data. One is used to determine the sensitivity threshold of SPM99 analysis in our clinical setting, and another was prepared for audit of HMPAO SPECT analysis techniques, by simulating pathological presentation of disease. We present the initial findings of the applications.

1. Introduction

Single photon emission computed tomography (SPECT) perfusion imaging of the brain is a useful tool in the diagnosis of dementia. There are currently several widely accepted semi-quantitative analysis techniques described in the literature that are useful in aiding in the interpretation of SPECT images. These methods offer an increase in the level of sensitivity and specificity over that of viewing alone. Validation of the techniques is usually based upon cross-reference with clinical observations and psychiatric tests, and lacks the inclusion of known ground truth. Basic validation has been performed on statistical parametric mapping (SPM) using simulated data sets [1], however the method used for the generation of validation data did not produce very realistic images. The use of simulated images allows analysis to proceed with the ground truth being known, giving a method of indexing the accuracy and precision of the analysis method. Full system simulation can be achieved through the use of Monte-Carlo codes. These methods offer simulation of all aspects of the imaging system on a photon by photon basis, but are limited by very long run times to simulate a useful number of detected events.

Another method of simulating gamma camera images is to make assumptions based on measured data, and incorporate this information into the simulation. One such method uses a measured or modelled point spread function (PSF) and proceeds to model the detected distribution of events from a single point source of activity, eliminating the need to simulate each photon. The image is then formed by the combination of PSF's for a given activity distribution. This method has been successfully implemented using different PSF models [2, 3, 4]. These methods can be extended to simulate abnormalities. We present a fast method for the simulation of realistic abnormalities in acquired normal data, incorporating an experimentally measured 3D PSF and existing normal reconstructed hexamethylpropyleneamine oxime (HMPAO) SPECT brain scans. The study reported here produces realistic lesions representing actual disease progression, using the Talairach atlas to synthesize the anatomical position of the abnormalities. Inter subject variability is achieved through the use of acquired normal subject images. Other methods have been described for the simulation of abnormal lesions for the validation of the performance of analysis, and various approaches have been used to simulate abnormal images [1, 5, 6].

2. Methods

2.1 Imaging, Processing and Analysis

Subjects were chosen retrospectively from the department database of carefully screened normal healthy controls, demonstrating no confounding problems. Subjects were injected, scanned and processed using the standard clinical procedure. The data is then reconstructed using filtered back projection (FBP) and extraneous facial activity is removed by interactive masking. A leave one out analysis was performed using SPM99

[1] Address for correspondence: MP26, Southampton General Hospital, Tremona Road, Southampton. SO16 6YD
Telephone: +44 (0)23 8079 5055 Email: tony.ward@suht.swest.nhs.uk

(Wellcome Department of Imaging Neuroscience, London, UK) to ensure the analysis method was not introducing false positives into the normal subjects, and consequently all subjects over 80 years of age were omitted from the data set.

SPM99 was used for image analysis as follows. Images were spatially normalised to the SPM SPECT template using a 12 parameter affine transformation and a 12 step iterative non-linear deformation (4 x 5 x 4 basis functions, bi-linear interpolation). The images were then smoothed with a 16mm FWHM 3D Gaussian kernel, resulting in matrix size of 91x109x91 voxels with an isotropic voxel size of 2mm. The count level was normalised to the average maximum counts on both sides of the cerebellum. Analysis was then performed by comparing a single individual subject to a group of control subjects on a voxel by voxel basis using a two-sample t-test ($p < 0.001$ uncorrected). Clusters were considered abnormal at $p<0.05$ (corrected).

A 200mm diameter, cylindrical water filled phantom was used to contain the point source of activity for measurement of the PSF. A 2.3MBq point source measuring approximately $1mm^3$ was placed at the centre of the phantom. The phantom was imaged at a distance of 15cm from the centre of rotation. The data is then processed using the same method as described above, and the measured PSF is scaled such that $\int PSF = 1$.

2.2 Simulation

VOI's were extracted for each of the regions defined by the Talairach Daemon (Research Imaging Centre, University of Texas, USA), which is a stereotactic brain atlas based upon the space defined by Talairach and Tournoux [7]. The daemon splits the Talairach atlas into 5 levels consisting of cerebral organisation, associated lobe, associated gyrus/sulcus, the type of matter and the Brodmann area. Each region was extracted into individual left and right-sided regions, which can be used to precisely position abnormalities in standard space, such that they can be used in the simulation system. The regions were then transformed into the standard space defined by the Montreal Neurological Institute (MNI), as used by SPM, using a transformation supplied by Matthew Brett [8].

Spatial registration is then performed on the image using SPM99 with a 12 parameter affine transformation and the resulting transformation is stored for future use. The VOI's to be simulated are selected and combined and/or masked with other VOI's to produce a pattern of abnormality with the intensity at each voxel representing the percentage perfusion reduction in that voxel. The stored transformation matrix for the image is reversed and applied to the standardised abnormality to transform it to the original subject space. Abnormalities can also be defined directly as discrete voxels or clusters of voxels in predefined geometric shapes, which are positioned interactively on the reconstructed subject image. The abnormality map is then convolved with the measured PSF to generate a simulated abnormality distribution. This image is inverted and multiplied by (as opposed to subtracted from) the normal image. This process preserves features of the original image, such as the natural image noise, at a relative level. The result is a simulated abnormal image with a known abnormality distribution. Equation 1 illustrates this concept mathematically for the generation of the simulated image I_{sim}.

$$I_{sim} = I_{sub} \times \left(1 - \left(I_{ab} \otimes I_{PSF}\right)\right)$$

Where I_{sub} is the subject image, I_{ab} is the image representing the abnormality contribution in subject space (relative to 1) and I_{PSF} is the measured and normalised point spread function.

2.3 SPM Sensitivity Assessment

This experiment was designed to ascertain the size and intensity detection threshold of SPM99 in the clinical setting. 21 subjects were selected retrospectively from the full normal database. From the group, a single control subject was used as a test subject and the remaining controls constitute the normal database for comparison using SPM. Four symmetrically paired abnormalities were stereotactically defined in the grey matter of the four major brain lobes in the reconstructed data. The abnormalities were initially defined in grey matter only, and the region size was increased by region growing and constraining the growth of the regions depending on the immediate neighbourhood of the voxels. Intensity was varied from 0% to 100% in 25% steps. Each abnormality was simulated, incorporated into the subject image and analysed. Table 1 gives a description of the abnormalities along with the stereotactic coordinates and the size of the abnormality. SPM analysis was performed using the standard protocol, and an abnormality was considered detected if significant regional hypoperfusion was evident.

Region	FR	FL	TR	TL	PR	PL	OR	OL
Origin	60 45 18	75 45 19	49 59 33	83 64 34	58 68 16	74 70 16	59 80 57	73 81 38
Size of region in voxels	152	158	146	159	149	151	144	145
	255	255	254	260	244	251	241	258
	340	358	359	344	347	348	353	348
	467	460	447	456	459	451	449	456

Table 1. *Description of abnormalities used in SPM99 sensitivity measurement, showing the origin of the region and the actual abnormality size in voxels.*
F = frontal lobe, T = temporal lobe, P = parietal lobe, O = occipital lobe. L= left, R = right side

2.4 HMPAO Brain SPECT Audit Data

The purpose of this experiment was to generate a set of abnormal subject images that could be used in the audit of cross centre analysis methods, facilitating a method of testing the efficacy of the methods. 29 normal subject images were available for this study. The studies were divided into two groups representing a normal database (group 1, n=19) and a group of subjects to be used for abnormality simulation to represent the unknown population (group 2, n=10). Age and sex variance was minimised within and across the two groups.

The abnormal group is subdivided to represent the four major dementia diseases, including Alzheimer's dementia (AD, n=2), vascular dementia (multiple infarct dementia (MID), n=2), fronto-temporal dementia (FTD, n=2), dementia with Lewy Bodies (DLB, n=2), and 2 unaltered subjects. The severity of the condition was randomised across the abnormal group. The data was simulated and assessed locally, with the help of a nuclear medicine consultant, using SPM to iteratively reach a realistic representation of the true disease. The data set can be split into 3 different types of abnormal image; 2 images representing control subjects with no abnormality, 6 images representing common dementia types and 2 MID images with varying sizes and weights of abnormality that offer a method of assessing the sensitivity of the analysis method under study. In subsequent sensitivity and specificity measurement, the MID patients are not included in the analysis results.

3. Results

Table 2 shows the number of significant detected abnormalities (non significant in brackets) against the size of the abnormality and the percentage reduction applied.

	150 Vx (5.8ml)	250 Vx (9.7ml)	350 Vx (13.5ml)	450 Vx (17.4ml)
25%	0 (0)	0 (0)	0 (0)	0 (0)
50%	0 (0)	0 (3)	0 (4)	1 (5)
75%	0 (2)	0 (5)	4 (8)	6 (8)
100%	0 (5)	4 (8)	8 (8)	8 (8)

Table 2. *SPM99 - Number of significant and non-significant detected abnormalities for varying size and intensity. Abnormality sizes are given in voxels and ml.*

The variation of sensitivity of abnormality detection with volume and intensity of perfusion deficit was determined. SPM analysis required a reduction of 50% for subsequent detection of significant abnormalities over the range of size studied, and similarly a minimum abnormality size of 9.7cc to detect any levels of significantly reduced perfusion. 100% sensitivity was only achieved for abnormalities larger than 13.5cc with a complete reduction in perfusion. Sensitivity was considerably higher for the temporal and occipital lobe abnormalities compared to the frontal and parietal lobes. No meaningful difference was witnessed between the two hemispheres of the brain.

The full audit data set has been analysed blindly using SPM99 by an experienced nuclear medicine physicist using our routine clinical protocol. The detection sensitivity and specificity for the introduced abnormalities were observed as 95% and 88% respectively for our clinical implementation of SPM99.

4. Discussion

This simulation method using measured PSF data and existing patient images offers fast simulation that cannot be achieved using more analytical approaches. The incorporation of the abnormality into the image through multiplication results in the preservation of the noise in the image at a relative level, removing the need to adapt

the abnormality to take this into account. The inclusion of a stereotactic atlas and image registration allows for the precise anatomical positioning of abnormalities, allowing the simulation of realistic disease presentation. Furthermore, as the images are based on real subjects, anatomical and physiological variation is already included in the images. Data generated using the simulation method could be used in the optimisation of different analysis methods for individual sites.

There are currently several limitation of this simulation system. This method produces reconstructed data and as such is not applicable to production of planar projection data such as would be required in the study of reconstruction methods. This may also require that other centres protocols would need to be adjusted to take this into account when performing quantitative analysis. The PSF used in this simulation is taken from an average distance inside a water filled phantom, intended to represent the PSF for the average voxel. Distance dependent factors affecting the PSF such as scatter and attenuation should be taken into account to provide a more realistic model of the PSF. While the effects of noise were minimised through the use of large activities for the PSF measurement, these effects cannot be totally removed, which adds uncertainty to the result. This could be minimised through the use of three-dimensional mathematical modelling of the PSF incorporating these parameters. Affine registration is used to transform the abnormality into the subject space, but does not take into account individual physiological differences, and errors may be evident in the positioning of the abnormalities. However, these errors would be generally small compared to the FWHM of the PSF, and are as such negligible.

The SPM99 sensitivity measurement allowed the diagnostic accuracy of the method to be assessed in our clinical implementation (table 2). By varying the size, intensity and position of the abnormalities, the threshold level could be obtained for each independent region. The method demonstrates the usefulness of the simulation system in generating data that can be used to assess analysis software. A set of data has been generated that will be used in the audit of cross centre analysis methods, and it is intended that this data will be used in a ^{99m}Tc HMPAO brain SPECT pilot audit study. Initial analysis indicates that all of the simulated images represent their intended presentation, and the results show variation in the detection of the introduced abnormalities.

5. Conclusion

We have presented a fast and realistic method of abnormality simulation in normal subjects, incorporating anatomical and physiological variation. The method offers the ability to generate stereotactic abnormalities of arbitrary or anatomical composition, allowing the pathology of disease to be simulated. The method has been implemented to test the sensitivity of our analysis methods, and highlights problems associated with it.

The method described is useful for the generation of data sets allowing for the optimisation of processing and analysis parameters, such that optimum sensitivity and specificity could be achieved for the method of choice through the interactive adjustment of these parameters.

6. References

1. E.A. Stamatakis, M.F. Glabus, D.J. Wyper, A. Barnes & J.T. Wilson. "Validation of statistical parametric mapping (SPM) in assessing cerebral lesions: A simulation study" *Neuroimage* **10**, pp.397-407, 1999.
2. J.S. Fleming & D.E. Simpson. "A technique for simulation of the point spread function of a gamma camera" *Phys. Med. Biol.* **39**, pp.1-17, 1994
3. Y. Ma, M.Kamber & A.C. Evans. "3D Simulation of PET Brain Images Using Segmented MRI Data and Positron Emission Characteristics" *Computerized Medical Imaging and Graphics* **17**, pp.365-371, 1993
4. H. Kim, B.R. Zeeberg, F.H. Fahey, A.N. Bice, E.J. Hoffman & R.C. Reba. "3D SPECT Simulations of a Complex 3D Mathematical Brain Model and Measurements of the 3D Physical Brain Phantom" *J. Nucl. Med.* **32**, pp.1923-30, 1991
5. S.J. Stapleton, C.B. Caldwell, C.L. Leonhardt, L.E. Ehrlich, S.E. Black & M.J. Yaffe. "Determination of Thresholds for Detection of Cerebellar Blood Flow Deficits in Brain SPECT Images" *J. Nucl. Med.***35**, pp.1547-55, 1994
6. C. Grova, P Jannin, A Biraben, I Buvat, H Benali, A.M. Bernard, J.M. Scarabin & B Gibaud. "A Methodology for Generating Normal and Pathalogical Brain Perfusion SPECT Images for Evaluation of MRI/SPECT Fusion Methods: Application in epilepsy" *Phys. Med. Biol.* **48**, pp.4023-43, 2003
7. Talairach J., Tournoux P. "Co-planar stereotactic atlas of the human brain" Thieme Medical Publishers, Inc. 1988
8. CBU Imaging web page - http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml

Three-Dimensional Voxel Morphometry of MR Brain Images Using Deformable Models, Relative Fuzzy Classification and Spatial Affinity

Runa Parveen[a1], Cliff Ruff[b], Darren McDonald[a], Tryphon Lambrou[a] and Andrew Todd-Pokropek[a]

Department of Medical Physics & Bioengineering, [a]University College London, [b]University College London Hospitals, UK

Abstract This paper presents an algorithm for classifying different tissue types in MR brain images in both normal and abnormal cases. Initially, an adaptive greedy contour model has been developed to separate the intracranial volume (ICV) from the scalp and skull. Next, to counteract the blurring of tissue boundaries due to the partial volume effect, an algorithm for fuzzy segmentation is presented which uses integrated fuzzy spatial affinity with statistical distributions of image intensities for each one of the three tissues – white matter (WM), grey matter (GM) and cerebrospinal fluid (CSF) respectively. Experimental results of brain tissue classification on 3D real and simulated MR data are provided.

1 Introduction

In human brain imaging and diagnosis, Magnetic Resonance Imaging (MRI) has become an important tool, providing good soft tissue contrast, its non-invasive nature is also an advantage. Extraction of the 3-D (three-dimensional) structure of soft-tissues is one of the active areas of research in computer-aided analysis and diagnosis. The classification of MR images of the brain, that is, assigning each voxel to a specific tissue type has received a considerable amount of attention. The soft tissue contrast enables complex and subtle brain structures to be clearly visualised. However, a unique correspondence of grey level ranges to different tissue types does not exist. It is difficult to detect anatomical structures due to unpredictable regions of interests (ROIs) in the presence of noise, finite spatial resolution and partial volume effects. The process of labelling voxels as members of different non-overlapping regions, whose union is the entire image, is called image segmentation [1]. In this paper, algorithms for three-dimensional segmentation of anatomical objects from MR brain images are presented.

Approaches to segmentation may be classified into different categories: edge-based detection represents the boundaries of objects where signal changes occur [2] while region-based methods [3,4] grow connected regions, which are homogeneous according to some measure of gray level or texture. Automatic thresholding and morphological operations have been used in many segmentation techniques [5,6]. Each of these groups may be further divided into subgroups: hard and fuzzy – depending on whether the defined voxels are described with a binary value of 1 or 0 (known as hard) or the amount of tissue belonging to the given tissue class of the voxel (known as fuzzy) [7-9]. However, a major disadvantage of the use of fuzzy techniques alone is that they do not incorporate information about spatial relationships. The limited spatial resolution of MR imaging and the complex shape of the tissue interfaces in the brain imply that a large portion of voxels are affected by partial volume (PV) voxels, i.e., voxels that contain a mixture of two or more tissue types, and not a single tissue.

In this paper, a deformable contour model has been used that is a modified version of the greedy algorithm [12] to delineate the ICV from the skull and scalp. After delineating the brain, we have used spatial affinity based relative fuzzy classification to label the regions of white matter, grey matter and CSF. In our experiment, we used T1-weighted MR brain images and experimental results of brain tissue classification on real and simulated brainweb 3D – MR data [14] are provided.

2 Methods

2.1 Deformable Contour Model

Active Contour Models (ACM) were first introduced by Kass et al [13] to deform a contour to extract a feature of interest within an image. This algorithm provides a promising framework for boundary detection through the solution of energy minimization using variational calculus. The energy function is an integral sum of different weighted energy constraints. Williams et al [12] modified the ACM with the greedy algorithm to make it more stable and flexible. The original contour model was attracted to contours with large image gradients and in the greedy algorithm, the location that gives the smallest energy value is chosen. In this study,

[1] Author Email: runa@medphys.ucl.ac.uk

a deformable model has been developed based on the modified version of the ACM [12] to delineate the volumetric brain structures.

The model requires a set of control points (*CPs*) to start with which roughly delineate the volume of interest on several slices. In this case, object boundaries do not need to be marked precisely and also it is not necessary to do this on every slice. These *CPs* are then regenerated by using cubic spline interpolation to produce the same number of evenly spaced *CPs* that are smooth in the first derivative and continuous in the second derivative, both within an interval and at its boundaries. These *CPs* are then sorted so that they are roughly adjacent to each other in the third dimension. Splines are then generated between each set of connecting *CPs* in the third dimension to generate a full set of *CPs*, i.e., one for each slice.

After generating the full set of *CPs*, the ACM is then used to adjust their position and shape to find the best nearby contour by optimizing the spline energies of the model. Each possible configuration of the contour is associated with the total energy dependent on the elasticity and rigidity of the spline. Among the energies, internal spline forces impose a smoothness constraint on the contour to control continuity, stretching and bending of a contour. In this case of elastic energy, the approximation of evenly spaced contour points has been maintained by calculating the average distance between all points for contours in the xy plane and in the z direction respectively, which guides the overall performance of the model. The external energy is a weighted combination of two terms: one attracts the model to edges and the other constrains the model to surround a homogeneous region. In our method, we move the *CPs* to the desired regions by normalizing the energy values for internal and external energies. Dependent on the direction of normal vector and the curvature of each connecting line between the *CPs* , we calculate the maximum gradient of internal energy (continuity and bending energy) I_d as well as for the external energies, E_d in the given search space s and the subscript d means the gradient. We also calculate the difference of energies, I_{ks} and E_{ks} with the minimum energy of each of these *CPs* respectively in the given s, where k is the number of neighbourhood points. Next, all the *CPs* are reparameterised by normalizing I_{ks} and E_{ks} with the maximum gradient I_d and E_d in ks space. New CPs are then constructed using the energy minimization to give the best fit to the structure of interest.

The three-dimensional deformable model is composed of a set of control points that are connected both in a two-dimensional plane and between slices. $v_{xy}(i)=\{x(i), y(i)\}$ and $v_{xyz}(i)=\{x(i), y(i), z(i)\}$ represent a section of contour in the xy plane and connected contours with components in xyz, where x, y and z are the spatial coordinates and i is the length along the contour. The total energy can be represented by:

$$E_{total} = \int \left[E_{dcm}\left[v_{xy}(i)\right] + E_{dcm}\left[v_{xyz}(i)\right]\right] di$$
$$where\ E_{dcm}\left[v_{xy}(i)\right] = \alpha E_{\text{int}}\left[v_{xy}(i)\right] + \beta E_{ext}\left[v_{xy}(i)\right]$$
$$and\ E_{dcm}\left[v_{xyz}(i)\right] = \alpha E_{\text{int}}\left[v_{xyz}(i)\right] + \beta E_{ext}\left[v_{xyz}(i)\right]$$

where, E_{int} represents the internal spline energy, E_{ext} represents the energy from image features. α and β are the weighting parameters to control the internal energy and the image energy, respectively. Figure 1(a) shows the results of applying the algorithm on simulated brainweb 3D-MR data [14].

2.2 Relative Fuzzy Classification with Spatial Affinity

The main aim in this study is to compensate for the blurring effect of tissue boundaries due to the partial volume effect. To overcome this effect, the statistical distributions of the image intensities of each tissue type p_{wm}, p_{gm}, and p_{csf} have been integrated with the fuzzy spatial affinity of each tissue element to classify the WM, GM and CSF. Classification assigns voxels/regions in an image into one or more specified classes and voxels can be a hard (binary) or fuzzy process. In the hard binary case, the voxels are assigned to classes with a value of 1 or 0, whereas in fuzzy classification, the process assigns the membership values for each tissue type in each voxel. Therefore, the membership values e.g. μ_{wm}, μ_{gm}, and μ_{csf} indicate the amount of tissue present in the voxel or the probability that that voxel belongs to the given tissue class. Alone the fuzzy process can not incorporate information about the spatial context making it sensitive to noise. In order to minimise this problem we calculate the connectedness between each group of tissues in the spatial domain.

The fuzzy classification process is designed to minimize the overall objective function F with respect to the membership functions μ_{ik} and the centroids v_k,

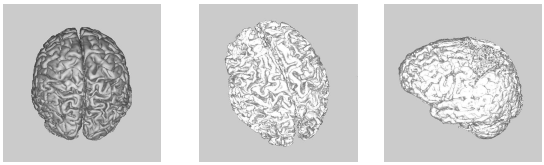
$$F = \sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik})^m \left(\| x_i - v_k \| \right)^2$$

where, x_i is the observation at voxel i, c is the number of clusters or classes, n is the image domain and the parameter m determines the amount of fuzziness of the resulting classification. The membership values of a voxel i to each class k are $\mu_{ik} \in [0,1]$ and $\sum_k \mu_{ik} = 1$ and the cluster centres v_k are upgraded using the membership values $v_k = \sum_{i=1}^{n} (\mu_{ik})^m x_i \Big/ \sum_{i=1}^{n} (\mu_{ik})^m$.

The spatial affinity between voxels has been taken into consideration in the classification process. Here, connectedness takes into account the adjacency of the voxels and the similarity of their intensity values. This affinity is then used to assign the strength of connectedness between the voxels and the tissue classes by using the strength of connectedness between the successive points in the connecting path. The final strength of connectivity within the same tissue class is the maximum strength along the path of minimum distance. The fuzzy affinity in a membership image is a function of (i) the fuzzy adjacency between the voxels, (ii) the homogeneity of the voxel intensities along the connecting path, (iii) the closeness of the voxel intensities and of the intensity-based features (mean and standard deviation) to some expected intensity and feature values for the tissue class and (iv) the relative location of the voxels. In general, the implementation of fuzzy connectedness for $c, d \in C: \mu_{fc}(c,d) = h(\mu_\alpha(c,d), f(c), f(d), c, d)$, where c, d are the image locations of the two voxels, $\mu_\alpha(c,d)$ is an adjacency function based on the distance of the two voxels, and $f(c)$ and $f(d)$ are the intensity of voxels c and d, respectively. Fuzzy connectedness is a fuzzy relationship in C, where $\mu_{fc}(c,d)$ is the strength of the strongest path between c and d, which has the smallest affinity along the path. In Udupa [10, 11], a hard binary relationship is used in C based on the fuzzy connectedness: $\mu_{fc}(c,d) = 1, if\ \mu_{fc}(c,d) \geq \theta \in [0,1]$, otherwise 0, and μ_α was chosen to a hard adjacency relation. In our study, a modification was used in the fuzzy connectedness algorithm to take into account spatial affinity relationship along with the fuzzy membership values. In this study, we have used an affinity relationship based on μ_{wm}, μ_{gm}, and μ_{csf}, the adjacency of the tissue elements, and the intensities without any extra weighting parameter. With this modification, we are able to assign the undefined voxels to a tissue class based upon their spatial affinity. Figure 1(b), and (c) show the classification result of WM, and GM respectively of simulated brainweb data [14] obtained from this fuzzy process.

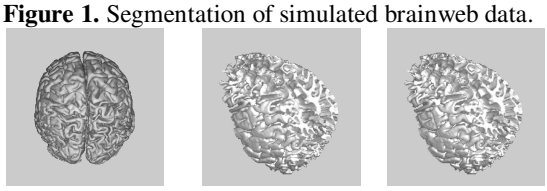
Results and Conclusion

This paper presents a new method for the segmentation of 3-D MR brain images. The proposed deformable contour model demonstrated here can successfully delineate the ICV from the scalp and skull. In this paper, a spatial affinity based relative fuzzy classification algorithm has been applied in the real and simulated 3-D brain images and the result on a healthy clinical data set is presented in Figure 2. Figure 3, shows the results of the proposed algorithm on the brainweb data, with Figure 4 showing the corresponding slices from the gold standard data for comparison. A comparison of tissue volumes is provided in Table 1. In comparison with previous published results, this new segmentation technique performs satisfactorily with regard to tissue classification of WM, GM, and CSF. This algorithm has been developed to identify and classify multiple sclerosis and tumour tissue and initial indications are that this classification technique can identify multiple sclerosis lesions.



1(a) ICV 1(b) WM 1(c) GM

Figure 1. Segmentation of simulated brainweb data.



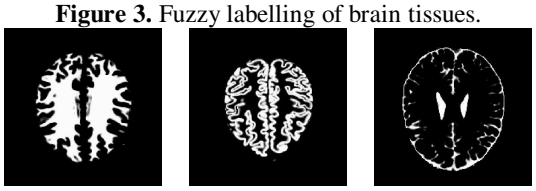
2(a) ICV 2(b) WM 2(c) GM

Figure 2. Result on a healthy clinical data set.



3a WM 3(b) GM 3(c) CSF

Figure 3. Fuzzy labelling of brain tissues.



4(a) WM 4(b) GM 4(c) CSF

Figure 4. Labelling obtained from gold standard data

	WM (voxels)	GM (voxels)	Brain (voxels)	CSF (voxels)	ICV = Brain + CSF
Simulated tissue volumes: (*BrainWeb*)	674777	902912	1577689	371945	1949634
Segmentation results on simulated data	694944	890611	1585555	384750	1970305
Error (%)	2.9%	1.4%	0.5%	3.4%	1.06%
Published Error results	3% [15]		4% [16]		4% [17]

Table 1: A comparison between the developed method and the simulated brain web data is presented for volume calculation of WM, GM, CSF and ICV with previous published result.

References

1. D.L. Pham, C. Xu & J.L. Prince. "A Survey of Current Methods in Medical Image Segmentation", *Ann. Rev. Biomed. Engineering,* 1998.
2. M. Bomans, K-H. Höhne, U. Tiede et. al.. "3-D Segmentation of MR Images of the Head for 3-D Display", *IEEE Trans. Medical Imaging*, 9(2), pp. 177-183, 1990.
3. R. Parveen & A. Todd-Pokropek. "Segmentation of MR Brain Images Using Region Growing Combined with an Active Contour Model", *Computer Aided Radiology and Surgery*, 2002.
4. H. Tang, E.X. Wu, Q.Y. Ma, et. al., "MRI Brain Image Segmentation by Multi-resolution Edge Detection and Region Selection", *Computerized Medical Imaging and Graphics*, 24, pp. 349-57, 2000.
5. M.E. Brummer e al, "Automatic Detection of Brain Contours in MRI Data Sets", *IEEE Trans. Medical Imaging*, 12(2): 153-166, 1993.
6. K.H. Hohne et al., "Interactive 3D Segmentation of MRI and CT Volumes Using Morphological Operations", *JCAT,* 16(2): 285-294, 1992.
7. J Bezdek et. al. "Review of MR Image Segmentation Techniques using Pattern Recognition", *Med. Phys.*, 20(4), pp. 1033-48, 1993.
8. L. P. Clarke et al., "MRI Segmentation: Methods and Applications", *Magn. Reson. Imag.*, 13(3), pp. 343-368, 1995.
9. J.C. Rajapakse et al. "Statistical Approach to Segmentation of single-channel cerebral MR Images",*IEEE TMI*, 16(2): 176-186, 1997.
10. J.K. Udupa et al, "Fuzzy Connectedness and Object Definition: Theory, Algorithms and Applications in Image Segmentation", *Graphical Models and Image Processing*, 58(3), pp. 246-261, 1996
11. J.K. Udupa et al, "Relative Fuzzy Connectedness and Object Definition: Theory, Algorithms and Applications in Image Segmentation", *IEEE Trans. PAMI*, 24(11): 1485-1500, 2002.
12. D.J. Williams et al, "A Fast Algorithm for Active Contours And Curvature Estimation", *CVGIP: IU*, 55(1), pp. 14-26, 1992.
13. M. Kass et al, "Snakes: Active Contour Models", *Proc. 1st Inter. Conf. Comp. Vision*, IEEE, pp. 259-268, 1987.
14. BrainWeb: Anatomical Model of Normal Brain, http://www.bic.mni.mcgill.ca/brainweb/anatomic_normal.html
15. Shattuck, DW et al. "Magnetic Resonance Image Tissue Classification Using a Partial volume model", *Neuroimage*, 13, 856-876, 2001.
16. Lemieux, L et al. "Fast, accurate and Reproducible Automatic Segmentation of the Brain in T1-weighted volume MRI data". *Magnetic Resonance in Medicine*, 42, 127-35, 1999.
17. Atkins, MS et al. "Fully Automatic Segmentation of the Brain in MRI", *IEEE Trans. MI*, 17(1), 98-107, 1998.

The effect of follicle volume measurement on clinical decisions

Mark J Gooding[a,*], Debbie Barber[b], Stephen Kennedy[b] and J Alison Noble[a]

[a]Medical Vision Laboratory, University of Oxford, UK
[b]Nuffield Dept. of Obstetrics and Gynaecology, University of Oxford, John Radcliffe Hospital, UK

1 Introduction

Currently, the growth of ovarian follicles under stimulation for fertility treatment is monitored using 2D ultrasound. The follicle size is recorded as the mean of two approximately orthogonal diameters across the follicle, in the plane in which the follicle appears largest. Human Chorionic Gondaotropin (hCG) is administered to the patient 36 hours prior to oocyte recovery, at a time when the oocytes are considered to be mature. The decision as to when to administer hCG is based solely on the mean diameter measurement, subject to blood estradiol levels being within acceptable limits.

In [1] we presented a semi-automatic method for the measurement of follicle volumes from freehand 3D ultrasound. This paper presents a study of the accuracy of the volume measurements, together with a study of the effect that using this volume measurement would have on management of the hCG decision.

2 Method

2.1 Accuracy of 3D measurement

In [1] we presented a semi-automatic method for the measurement of follicle volumes from freehand 3D ultrasound. This method uses a level set based region growing algorithm to segment manually identified follicles in 3D, whilst simultaneously interpolating the surface where data is absent. To assess the accuracy of this 3D volume measurement, volumes measurements from the reconstruction system were compared to aspirated follicle volumes as follows; The ovaries from 9 patients undergoing IVF treatment were scanned prior to oocyte recovery. The entire ovary was scanned for the 3D scan and all follicles were reconstructed. To aid identification of each follicle at the time of aspiration the clinician who would perform the follicle aspiration observed the 3D scan and labelled diagrams were drawn for each scan by the scanning clinician. Still images of the significant follicles were also printed and labelled. Although ideally all follicles would be measured in each ovary, if an ovary contained more than about 6 follicles, identification of follicles at aspiration is very difficult. Therefore only the follicles for which the aspirating clinician could be confident of a correct identification had volumes measured and recorded. Aspirated volumes were recorded to the nearest 0.5ml. The current clinical measure of follicle size recorded is the mean diameter. This assumes that the follicle is spherical and therefore that the diameter is representative of the follicle volume. To test this assumption, mean diameter measurements were also made at the time of the 3D scan. The volumes of the follicles were estimated from the clinical diameter measurements using a spherical model, for comparison to the aspirated volume.

2.2 Effect of volume on clinical decision

For a new system to be adopted for use in a clinical context it is necessary to show that quality of treatment management is at least equivalent to that of the current system in use. For fertility treatment management the quality of treatment management must be assessed for the decision as to when to administer hCG.

After a period of about 1 week of monitored hormonal stimulation of the ovaries, the follicles are nearing maturity. At this stage the decision as to when to aspirate the follicles must be made. The hormone hCG must be administered 36 hours before aspiration, to complete the maturation. Currently the decision of when to administer hCG and aspirate the follicles is made largely on an assessment of the number of mature follicles. A follicle is considered mature if it is greater than 18mm in mean diameter, and is considered post-mature if greater than 25mm. The hCG decision must be generally consistent between systems, if the reconstruction system is to be used for treatment management. It must be shown that the same decision is made whether based on volume measurement or based on mean diameter. In cases of a difference of decision, it is necessary that the volume measurement provides a better tool for treatment management.

*gooding@robots.ox.ac.uk

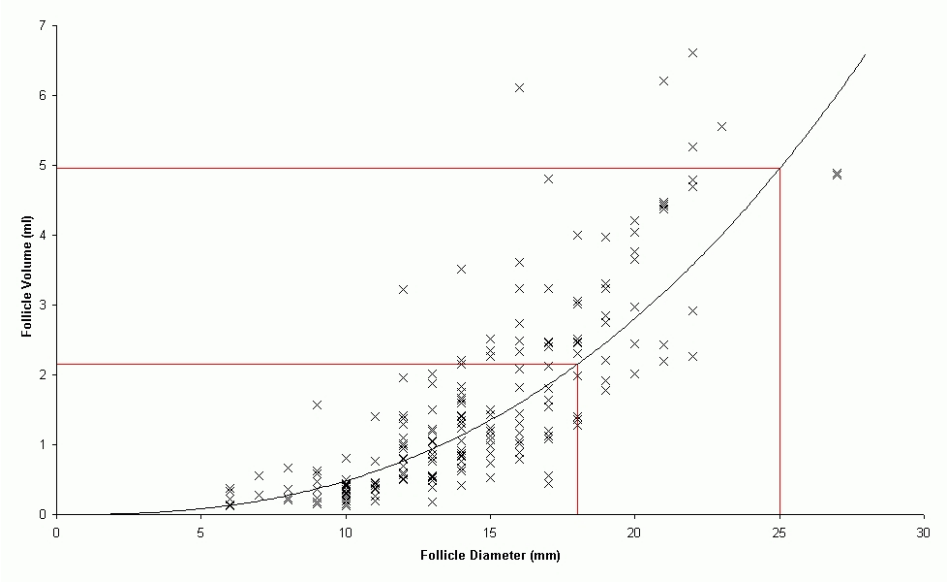


Figure 1. This figure shows the volume estimated from 3D reconstruction plotted against the clinicians diameter measurement. The board relationship established aids the clinician in understanding the new volume measurement in terms of the familiar diameter measurement. 2.14ml is equivalent to 18mm and 4.77ml is equivalent to 25mm.

To investigate whether the use of volume instead of diameter affects this decision, a simulation of this decision was used. Both diameter and volume measurements from 16 patients were collected. These measurements were presented to a clinician independently and in a random order. The clinician was asked to make a decision as to whether the follicles were mature and if the patient was ready for the hCG injection. The clinician was given access to the general relationship between volume and diameter to assist them in understanding the volume measurements. The measurements from all 16 patients were included randomly twice, to evaluate if the clinician's decision was consistent. The 32 diameter and 32 volume sets were coded so that the clinician was not aware of the correspondence between diameter data sets and volume data sets, or between repeated measurements. However since the clinician was not accustomed to making an hCG decision based on volume, it was necessary to establish a parametric relationship between mean diameter and volume. This relationship was found as described in [2], by measuring the volume and diameter of 191 follicles and fitting a curve through the data as the model of the realtionship. The clinician was allowed to refer to this relationship, as presented in figure 1, when presented with the volume data, such that an equivalent diameter could be found.

3 Results

3.1 Accuracy of 3D measurement

The data acquired consisted of 21 follicles for which aspirated volume, reconstructed volume, and diameter measurements were available. The low number of follicles in this part of the study is a result of the difficulty in identifying follicles at aspiration once the first follicle has been aspirated. This problem was also identified in [3], where only the largest follicle was measured in each of 25 patients.

Use of the Bland-Altman plot [4], shown in figure 2, illustrates that the estimated volume underestimates the aspirated volume by about 25%. The reasons for the underestimate have not been identified, and it is noted that the clinical measurement also underestimates the aspirated volume by about 25%. However, our main concern is with the reproducibility compared to the clinical measurement, since a measurement with a proportional error can still be used as an effective measure, but a measurement with high variance cannot.

Qualitatively it can be seen that the measurement made by the 3D system has a lower spread than the clinical data. Quantitatively, the variance of each data set from a linear fit to the data was $1.13ml^2$ and $0.43ml^2$ for the clinical and automated data sets respectively. For an F-test to be used to determine if the difference in variance between these sets is significant, it is necessary that the sets follow a Normal distribution. A chi-squared test was performed to check for normality. None of the data sets were rejected from being considered Normal. The F-test shows that the 3D reconstruction system has a significantly different variance to both the clinical measurements (p=0.966).

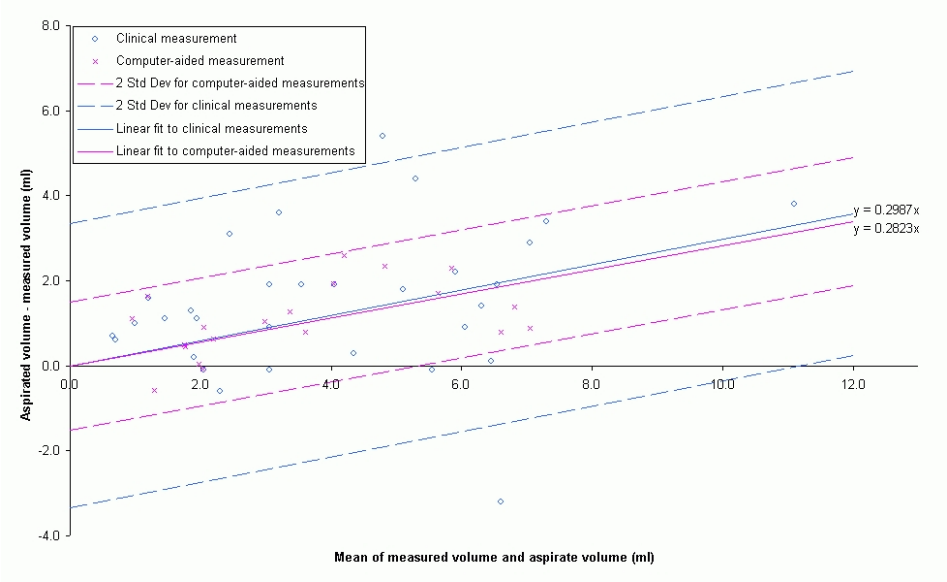


Figure 2. This graph shows the Bland-Altman plot for the reconstruction volume measurements and clinical volume estimates. Both the reconstruction method and clinical measurement have similar bias. The reconstruction method has a lower variance that the clinical measurement.

This variance is much lower than that of the clinical measurement, and therefore the 3D reconstruction system can be considered to be out-performing the current clinical measurement method.

3.2 Effect of volume on clinical decision

Generally the hCG decision made by the clinician was consistent for both presentations of each set of patient data, whether based on diameter and volume. For both diameters and volumes measurements there was only one patient (6.25%) for whom the decision of the clinician was different on the second occasion. However, this discrepancy occurred with a different patient for each measurement method. This is not unexpected since borderline cases will exist independent of the measurement method used. This experiment does show that the decision made by the clinician is highly repeatable.

For two patients (12.5%) the decision of the clinician differed when using volume measurements rather than diameter measurements. For these patients it is necessary to examine the reason for the discrepancy. In both cases the volume measurements were larger than would be expected given the diameter measurements, leading the clinician to decide on the basis of volume that hCG should be administered, whereas on the basis of diameter hCG would be postponed. Figure 3 shows the 3D reconstruction for an ovary from one of these patients, together with the plane in which the diameter was measured manually. The foreground follicle corresponds to the follicle for which the diameter measurements are shown. In the reconstruction is can be seen that the plane of measurement was perpendicular to the major axis of the follicle and as such the diameter measurement can be considered to be in error. In a protocol where the aim is to find the plane of maximum diameter it is expected that any error in finding this plane will lead to underestimation of the follicle size, leading to the hCG injection being incorrectly postponed.

4 Conclusion

It has be shown that the automated volumetric follicle reconstruction method presented in [1] has significantly lower variance than the state-of-the-art 2D clinical measurement. It should be expected that the diameter measurement would have higher variation because the diameter measurement cannot account for the shape of the follicles as they deviate from the spherical model. Despite this variation, there existed a large degree of agreement between hCG decision based on diameter measurement and those based on volume measurement. In those cases where there was a difference in the decision made, it is apparent that the volume measurement is of assistance, because the true size of the follicle is known. Therefore use of this measure may lead to better management of the treatment. However there still exists borderline decision cases even when using volume measurement. This should be expected because the decision is still fairly subjectively based on the size of the follicle and the expected oocyte yield. Use of expert systems [5] may be of benefit in such a well specified decision environment. Although it was

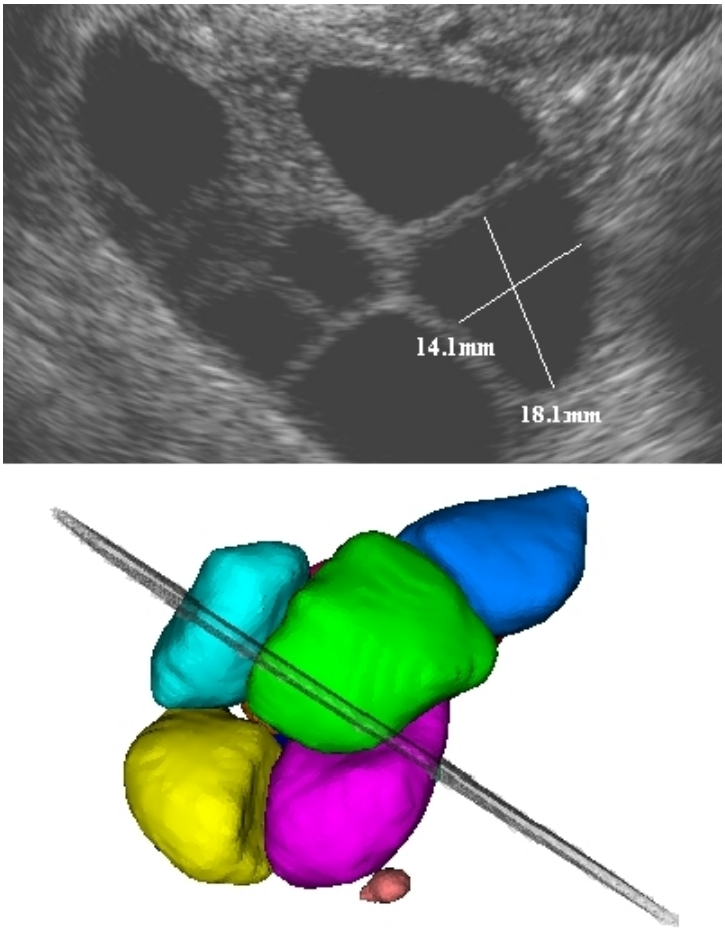


Figure 3. This figure shows the plane in which the clinician made the diameter measurement, together with the 3D reconstruction of the follicles within the ovary. This plane was perpendicular to the major axis of the follicle show with measurements. The equivalent follicle can be seen in the foreground on the reconstruction. In such cases the diameter measurement can be seen to be in error.

possible for the clinician to base decisions on volume measurements, there was a need to have the relationship with diameter available to help the clinician understand how to interpret the volume measurement. This need will persist to some degree when/if volume measurements are adopted clinically.

Further work should consider the effect on treatment outcome of management based on follicular volume. Variability of decision between clinicians should also be considered.

Acknowledgements

Mark Gooding is funded by the EPRSC as part of the MIAS-IRC. (GR/N14248).

References

1. M. J. Gooding, S. H. Kennedy & J. A. Noble. "Volume reconstruction from sparse 3D ultrasonography." In *Proceedings of MIUA*, pp. 205–208. 2003.
2. M. J. Gooding. *Volume reconstruction and measurement of ovarian follicles from freehand 3D ultrasonography*. Ph.D. thesis, Department of Engineering Science, University of Oxford, 2004. (Submitted).
3. A. Kyei-Mensah, J. Zaidi, R. Pittrof et al. "Transvaginal three-dimensional ultrasound: accuracy of follicular volume measurements." *Fertility and Sterility* **65(2)**, pp. 371–376, 1996.
4. J. Bland & D.G.Altman. "Statistical methods for assessing agreement between two methods of clinical measurement." *The Lancet* **i**, pp. 307–310, 1986.
5. I. Jurisica, J. Mylopoulos, J. Glasgow et al. "Case-based reasoning in ivf: Prediction and knowledge mining." *Artificial Intelligence in Medicine* **12(1)**, pp. 1–24, 1998.

Image-based Patlak analysis of dynamic ^{18}FDG PET studies: incorporating recovery-loss correction and blood curve normalisation to improve quantitative accuracy

Heather A Williams[a,c], Peter J Julyan[a], Malcolm Ranson[b], Jamal Zweit[c], David L Hastings[a]

[a]North Western Medical Physics, Christie Hospital NHS Trust, Manchester, UK
[b]Department of Medical Oncology, Christie Hospital NHS Trust, Manchester, UK
[c]Radiochemical Targeting and Imaging Group, Paterson Institute for Cancer Research, Manchester, UK

Abstract. Tumour glucose metabolism can be quantified *in vivo* using ^{18}F-fluoro-deoxy-glucose positron emission tomography (FDG PET) and Patlak graphical analysis, using image-derived input functions for the tumour and blood. However, the accuracy of uptake measures obtained by this method is undermined by effects arising from the limited spatial sampling of PET imaging devices. This paper describes the development and clinical validation of techniques to minimise and correct for these effects: optimised ROI analysis, concentration recovery correction and blood activity-time curve (BATC) normalisation.

1 Introduction

Positron emission tomography is a functional imaging technique, a sophistication of the radioisotope imaging techniques used in conventional nuclear medicine, which maps the biodistribution of a radiopharmaceutical incorporating a positron emitting isotope through coincident detection of the gamma photons produced when the emitted positrons annihilate with electrons within tissue. Coincident detection of annihilation photon pairs affords PET superior imaging performance compared to conventional techniques, together with the ability to quantify tracer concentration *in vivo*. Measures of radiopharmaceutical concentration are typically derived from PET images using region of interest (ROI) analysis. Uptake is most commonly expressed in terms of the radioactivity concentration at a fixed time post-injection, normalised to appropriate factors such as injected radioactivity and patient mass. ROI analysis can also be applied to a dynamic series of PET images to obtain blood and tissue activity-time curves which can act as input functions for kinetic models, enabling quantification of uptake in terms of the biological parameters associated with the uptake mechanism. This image-based approach avoids the risks associated with rapid arterial sampling, the traditional method of obtaining the BATC.

The majority of PET studies use the radiopharmaceutical ^{18}F-fluoro-deoxy-glucose (FDG) to detect and characterise tumours by means of their elevated glucose metabolism. In addition, pre- and post-therapy scans can be used to identify residual disease and assess response to treatment. The change in FDG uptake measures between pre- and post-therapy scans could be used to classify response [1] using a system analogous to that currently employed in anatomical imaging [2].

As PET imaging devices are limited by finite spatial sampling and statistical noise, the radioactivity distribution seen in the PET image is not an exact representation of the true distribution within the imaged object. For a PET scanner operating as a 2D slice imaging device, the effect of finite spatial sampling can be described in terms of recovery loss and spillover in the transaxial plane, and the partial volume effect in the axial direction. Recovery loss and spillover arise from the convolution of the true radioactivity distribution with the scanner response function, recovery loss being the reduction in the contribution from radioactivity in the lesion to a ROI-based concentration measurement and spillover being the contribution from radioactivity in surrounding tissues. The partial volume effect arises from signal averaging within the volume represented by the ROI pixels. For FDG-avid tumours, this is most pronounced for the first and last PET image slices in which the tumour appears, as these correspond to volumes partially occupied by normal tissue containing a much lower ^{18}F concentration.

The impact of limited spatial sampling on measures of tumour concentration measures can be reduced by careful ROI analysis, but there remains a discrepancy between measured and true concentration for smaller structures. If uncorrected, this discrepancy would lead to inaccurate measures of radioactivity concentration, which could in turn undermine the accuracy of response assessment via FDG PET. The aim of this work was to improve the quantitative accuracy of FDG uptake measures acquired by Patlak graphical analysis, using tumour activity-time curves (TATCs) obtained by optimised ROI analysis and concentration recovery correction and BATCs normalised to the venous ^{18}F concentration.

Correspondence should be addressed to Heather A Williams, e-mail: HeatherA.Williams@cmmc.nhs.uk

2 Methods

2.1 Imaging and blood sampling protocol

Five patients with stage IIIB or IV primary non-small cell lung cancer (NSCLC) underwent dynamic PET imaging, initiated to coincide with intravenous administration of 200MBq ^{18}FDG. The dynamic imaging protocol consisted of 18 frames, those at the start of the series being shorter in order to adequately characterise the BATC peak [3]. A transmission scan was performed prior to administration, to confirm correct positioning and facilitate attenuation correction. 3 ml venous blood samples were taken from the arm contralateral to the injection site at approximately 45, 55 and 65 minutes after administration. For the second patient, the sampling line became blocked and samples were withdrawn from the injection site, after first withdrawing and discarding 10ml of blood [4].

Each frame was reconstructed by filtered back-projection (FBP) using a Hanning filter with a cut-off frequency of 0.156 mm^{-1} (67% Nyquist), and measured attenuation correction (MAC). In addition, the frame in which radioactivity concentration peaked within the aorta was reconstructed using ordered subsets expectation maximisation (OSEM) with 28 subsets, 2 iterations and a 3 mm^{-1} Gaussian post-filter, with segmented attenuation correction (SAC). The radioactivity concentrations in the venous blood samples were determined using an automatic sample counter calibrated for ^{18}F, and decay-corrected to the time of administration.

2.2 TATC definition

Tumour ROIs were defined on the 18th frame, acquired from 60 to 75 minutes post-injection. Four different ROI analyses were used: conventional, focussed, and optimised ROI analysis with and without recovery loss. Conventional and focussed analyses reflect the methods most commonly used to obtain radioactivity concentration measures. Conventional analysis defines radioactivity concentration as the mean of the average concentrations within each of a series of ROIs drawn by eye around the outer edge of the lesion on each of the transaxial slices in which the lesion was visible. Focussed analysis was also carried out for cavitated lesions, by applying conventional analysis to the most FDG-avid focus within the lesion. Conventional and focused ROIs were copied onto all preceding frames to produce $TATC_{conc}$ and $TATC_{foc}$, respectively.

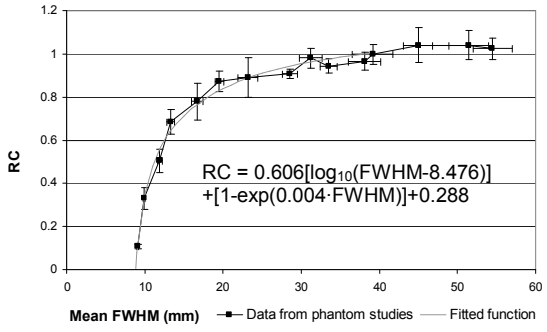


Figure 1. Variation of RC with FWHM for optimised ROI analysis.

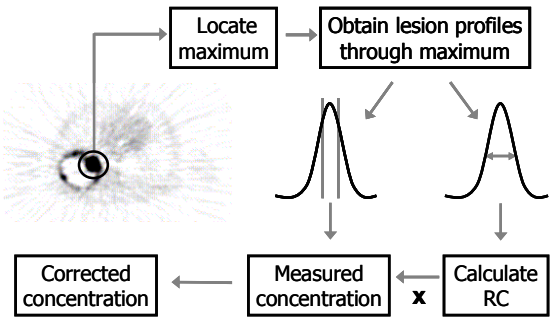


Figure 2. Concentration recovery correction process

Optimised ROI analysis defines radioactivity concentration as the mean of the average concentrations within each of a series of circular ROIs of diameter equal to the clinical spatial resolution (8.4 mm for ^{18}F), centred semi-automatically over the peak activity within the lesion on all but the first and last transaxial slices in which the lesion was visible. In a previous comparison of several ROI analyses, this was found to minimise the discrepancy between measured and true radioactivity concentration [5]. This discrepancy is quantified using the recovery coefficient (RC), defined as the measured radioactivity concentration divided by the true radioactivity concentration, in the absence of surrounding activity. Figure 1 shows the variation of RC with the full width at half-maximum (FWHM) of the lesion activity profile, which formed the basis of the concentration recovery correction process shown in Figure 2. Phantom studies have shown that this process produces corrected ^{18}F radioactivity concentrations within 10% of the true concentration for lesion diameters greater than 14.3mm and lesion to surrounding radioactivity ratios less than 3.4:1 [5]. Optimised ROI analysis was applied to the tumour images from frames 12 to 18, both with and without concentration recovery correction, to produce $TATC_{opt}$ and

$TATC_{optcorr}$ respectively. Contrast in earlier frames was inadequate for accurate central placement of ROI and accurate determination of FWHM.

2.3 BATC definition and normalisation

The blood ROIs were defined over the aorta, as BATCs obtained from this vessel have been shown to most closely resemble the arterial radioactivity-time curve [6]. Conventional ROI analysis was applied to those image slices where the aorta appeared perpendicular to the transaxial plane [7], from the OSEM-reconstructed frame in which radioactivity concentration peaked within this structure. These were then copied to all preceding frames to create BATC.

The BATC was matched to the venous radioactivity concentration by subtracting the mean difference between the BATC and the venous sample concentrations at 45, 55 and 65 minutes post-injection from all BATC data-points, creating $BATC_{norm}$. Where actual blood sample times deviated from the scheduled sample times, the venous radioactivity at the schedule times was deduced by linearly interpolating between, or extrapolating from, the existing data. This method is based on the work of Hoekstra and colleagues [8].

2.5 Kinetic analysis

Kinetic analysis was performed using the Patlak graphical method [9], adapted to account for the discrete nature of PET data. In this model, TATC and BATC data are related by Equation 1:

$$NC(n) = K_i \cdot NT(n) + V_d \quad (1)$$

where $NC(n) = \dfrac{C_T(n)}{C_P(n)}$ = normalised counts

K_i = rate constant for unidirectional, steady state trapping of ^{18}FDG into the tumour, min^{-1}

$NT(n) = \dfrac{\sum_0^n [C_P(n) \cdot \Delta t(n)]}{C_P(n)}$ = normalised time

V_d = volume of distribution for unmetabolised FDG (approximate), unitless

C_T = FDG concentration in tumour, $kBq \cdot ml^{-1}$

C_P = FDG concentration in blood, $kBq \cdot ml^{-1}$

n = Dynamic frame number

Δt = Dynamic frame duration, minutes

Plots of NC against NT were obtained for each of the four TATCs with BATC, and $TATC_{optcorr}$ with $BATC_{norm}$. Values of K_i were obtained from a least-squares fit to the linear phase of the curve, corresponding to frames 12 to 18, and the correlation between NC and NT quantified using Pearson's coefficient (r).The significance of changes in K_i with different input functions was assessed at the $p = 0.05$ level using analysis of covariance and Fisher's (F) test.

2 Results

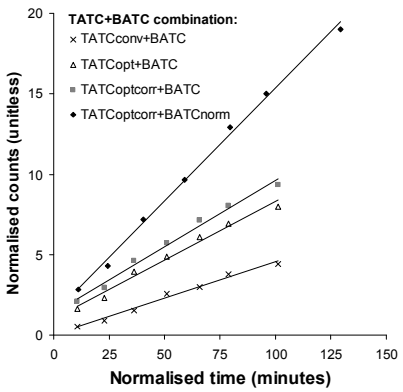


Figure 3. Patlak plots for patient 1, for various TATC+BATC combinations.

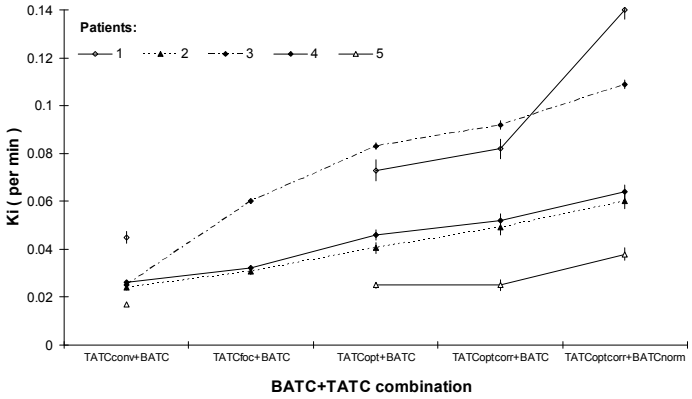


Figure 4. Variation in K_i with BATC+TATC combination for all five patients. Error bars denote the standard error in K_i

Good fits were obtained to the linear phase of all plots ($r \geq 0.98$). The differences between K_i for $TATC_{conv}$+BATC and $TATC_{foc}$+BATC were statistically significant from those for $TATC_{optcorr}$+BATC in all cases. For patients 3 and 5, the difference in K_i was also significant between plots for $TATC_{opt}$+BATC and $TATC_{optcorr}$+BATC. The differences between K_i for $TATC_{optcorr}$+BATC and $TATC_{optcorr}$+$BATC_{norm}$ were statistically significant in all cases.

4 Discussion

The imaging methodology was well-tolerated by patients with advanced NSCLC. Optimal ROI analysis, concentration recovery correction and BATC normalisation were successfully applied and generated highly-correlated Patlak plots from which the kinetic parameter K_i was obtained. Figures 3 and 4 show a progressive increase in K_i from conventional to focussed, then focussed to optimised ROI analysis can be attributed to the ROI becoming smaller, and encompassing a more homogeneous population of pixel values which are more representative of the true radioactivity concentration in tumour. An additional increase in K_i is observed with recovery loss correction for 4 of the 5 patients, although this is only statistically significant for 2 patients. K_i increased still further with BATC normalisation. Neither concentration recovery correction nor BATC normalisation degraded the quality of the linear fit.

5 Conclusion

This method is clinically practicable and successfully corrects for the effects of limited spatial sampling, preventing underestimation of K_i and potentially improving the accuracy of response assessments using this index of FDG uptake.

References

1. H. Young, R. Baum, U. Cremerius et al "Measurement of Clinical and Subclinical Tumour Response Using [^{18}F]-fluorodeoxyglucose and Positron Emission Tomography: Review and 1999 EORTC Recommendations", Eur J Cancer **35** pp 1773-1782, 1999.
2. P Therasse, S.G. Arbuck, E.A. Eisenhauer et al "New Guidelines to Evaluate the Response to Treatment in Solid Tumours", J Natl Cancer Inst **92** pp 205-216, 2000.
3. H. Minn, K.R. Zasadny, L.E. Quint et al "Lung Cancer: Reproducibility of Quantitative measurements for Evaluating 2-[F-18]-Fluoro-2-deoxyglucose Uptake at PET", Radiology **1969** pp 167-173, 1995.
4. C.J. Hoekstra & A.A. Lammertsma "On the use of the injection catheter for venous blood sampling in quantitative FDG PET studies", Eur J Nucl Med **27** pp 1579, 2000.
5. H.A. Williams & P.J. Julyan "What you see aint what you get: investigating concentration recovery in PET", Nuc Med Comm **24** pp 456, 2003.
6. A.P. van der Weerdt, L.J. Klein, R. Boellard et al "Image-Derived Input Functions for Determination of MRGlu in Cardiac ^{18}F-FDG PET Scans", J Nucl Med **42** pp 1622-1629, 2001.
7. G. Germano, B.C. Chen, S. Huang et al "Use of the Abdominal Aorta for Arterial Input Function Determination in Hepatic and Renal PET Studies", J Nucl Med **33**: pp 613-620, 1992.
8. C.J. Hoekstra, O.S.Hoekstra & A.A. Lammertsma "On the use of image-derived input functions in oncological fluorine-18 fluorodeoxyglucose positron emission tomography studies", Eur J Nucl Med **26** pp 1489-1492, 1999.
9. C.S. Patlak, R.G. Blasberg & J.D. Fenstermacher "Graphical Evaluation of Blood-to-Brain Transfer Constants from Multiple-Time Uptake Data" J Cereb Blood Flow Metab **3** pp 1-7, 1983.

Automated assessment of retinal image field of view

Alan Fleming[a], Sam Philip[a] , Keith Goatman[a], John Olson[b], Peter Sharp[a]

[a]Department of Biomedical Physics and Bioengineering, University of Aberdeen,
[b]Grampian Diabetes Retinal Screening Programme, Woolmanhill Hospital, Aberdeen

Abstract. Screening programmes using retinal photography for the detection of diabetic eye disease are being introduced in the UK. The large number of images involved has encouraged a number of groups to develop software to detect the earliest signs of disease. However, for such analysis to be effective one must first ensure that image quality meets two standards: that the image has sufficient clarity, and that the correct region of the retina has been photographed. This paper describes a solution to the latter problem based on the detection of known features, including the fovea and optic disc, in the retinal images. Constraints on the location and extent of the detected features have been devised that allow the image field of view to be verified. Results are presented for 1067 images from the Grampian diabetic retinal screening programme.

1. Introduction

Diabetic retinopathy (DR) is the major cause of blindness among the working-age population in the UK. In the majority of cases, the blindness is preventable if it is detected early enough. However, in its early stages the disease is entirely asymptomatic, which has necessitated the development of nationwide screening programmes for diabetic eye disease. As screening programmes require the examination of large numbers of images, software to automatically detect whether disease is present has been developed by a number of groups [1,2]. For this analysis to be effective, one must first ensure that image quality meets certain standards. The first is that the image clarity must be adequate to view relevant lesions if they are present [3]. The second is that the photograph should include the correct region of the retina, approximately centred on the fovea. No work has previously been published which describes automatic assessment of this second important aspect of image quality.

This paper describes a fully automated technique that assesses the adequacy of FOV of images in a DR screening programme. The approach is based on the accurate location of the essential features in the retinal image. Metrics are defined using the relative locations and sizes of these features which must satisfy certain conditions for the image to be accepted as having an adequate FOV.

2. Methodology

2.1. Source and type of images

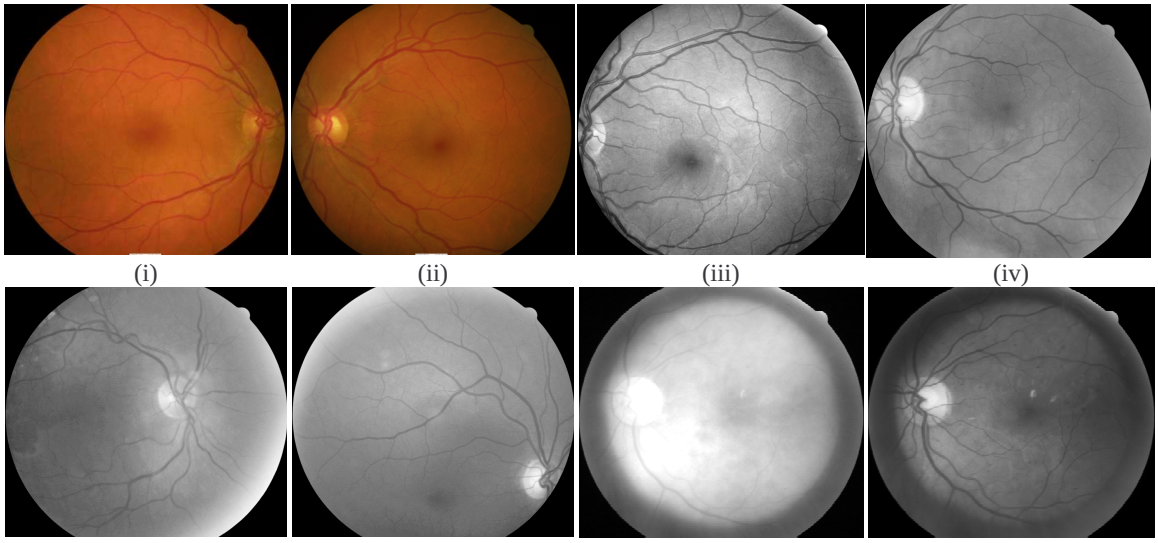
The images for the study are taken from patients attending the Grampian Diabetes Retinal Screening Programme. A 45° single field disc/macula photograph was taken of each eye. Images are acquired using a Canon D30 digital camera attached to either a Canon CR5-45NM or a Canon CR6-45NM fundus camera. Image size is 2160 by 1440 pixels with the fundus occupying a cropped disc inside a region of 1600 by 1440 pixels. The training set for the development of automated methods contained 395 images. The test set contained 1067 images, representing one month's images.

The adequacy of the image FOV was graded by a clinical research fellow using the criteria listed in Table 1, with examples shown in Figure 1. All grading was performed on a 22" monitor (Iiyama VM Pro 514) displaying the green plane of the colour image, which has the greatest feature contrast. The locations of the fovea and optic disc (OD) and which eye (left or right) were also recorded for comparison with the automated results.

Criterion	Classification	Requirement for adequate FOV
Optic disc complete	True / false	True
Macula complete	True / false	True
Temporal arcades complete	True / false	True
Small pupil artefact present	True / false	False

Table 1. Grading data recorded by the clinical research fellow.

Corresponding author: Alan Fleming (a.fleming@biomed.abdn.ac.uk)



(i) (ii) (iii) (iv)

(v) (vi) (vii) (viii)

Figure 1. Examples of retinal images: (i) adequate FOV, right eye, (ii) adequate FOV, left eye. The remaining images have inadequate FOV: (iii) OD incomplete, (iv) superior arcade incomplete, (v) macula incomplete, (vi) inferior arcade absent, (vii) & (viii) small pupil artefact present, red and green image planes.

3. Automated feature detection methods

3.1. Vessel detection

The cropped circular region occupied by the fundus image can be selected by intensity thresholding followed by morphological operations to clean the result. The image intensity is then scaled so that the maximum intensity is 255. The first stage of vessel detection is to create image g in which a pixel is given a positive value if an edge filter gives a positive response on one side and a negative response on the other, at a distance from the pixel of the order of half the expected vessel width. More precisely this can be written:

$$g(x,y) = \min(e(x_1,y_1), -e(x_2,y_2))$$

where e is the response of the edge filter, $(x_1,y_1) = (x,y) + (u,v)$, $(x_2,y_2) = (x,y) - (u,v)$ and (u,v) is a vector whose length is of the order of half the expected vessel width in a direction perpendicular to the vessel axis, (Figure 2(i)). Next morphological erosion with a linear structuring element perpendicular to the edge filter is applied. The entire vessel detection is applied at 18 orientations and the maximum, denoted G, is taken of the 18 results.

A constant threshold applied to G is not suitable due to variability between and across images. Instead, a range of thresholds coupled with basic shape analysis has proved more effective. A sequence of four thresholds is generated from the histogram of G and applied to create four binary images. Each of the binary images is dilated to remove isolated empty pixels, skeletonised, and has spurs removed. The skeleton is broken at junctions to create two-ended segments. The straightness of each segment is calculated by taking the ratio of the Euclidean distance between its endpoints and its arc length. Segments with length less than 90 pixels and straightness greater than 1.4 are rejected. Segments which pass these size and straightness criteria from any of the four binary images are combined. An example segmentation result is shown in Figure 2(ii): the vector (u,v) was deliberately chosen to ensure that the widest vessels are detected since the goal is to detect the temporal arcades.

3.2. Temporal arcade detection

The temporal arcades are the main veins and arteries that originate from the OD and extend in the direction of the temples: they have a roughly elliptic form with one apex on the OD, and are roughly symmetrical about the fovea. Since only approximate locations are required at this stage, the vessel image can be sub-sampled at 1:32 ratio by summing pixels in 32×32 blocks, to increase processing speed. The elliptical shape is detected using a Hough transform repeated for a set of 90 left- and right-hand semi-ellipses at various inclinations, aspect ratios and sizes, (Figure 2(ii)). The maximum response in Hough space gives the semi-ellipse location, inclination, size, aspect ratio and whether it is right- or left-handed. The set of pixels that contributed to this point in Hough space are taken to represent the main temporal arcades. Their lengths are calculated by finding the pixels which

contributed to the maximum point in Hough space. These pixels are divided into two sets using a line drawn through the OD and fovea (locations determined as below). These sets of pixels are skeletonised and then the arc length of each is calculated to give an estimate of the length of the superior and the inferior temporal arcades.

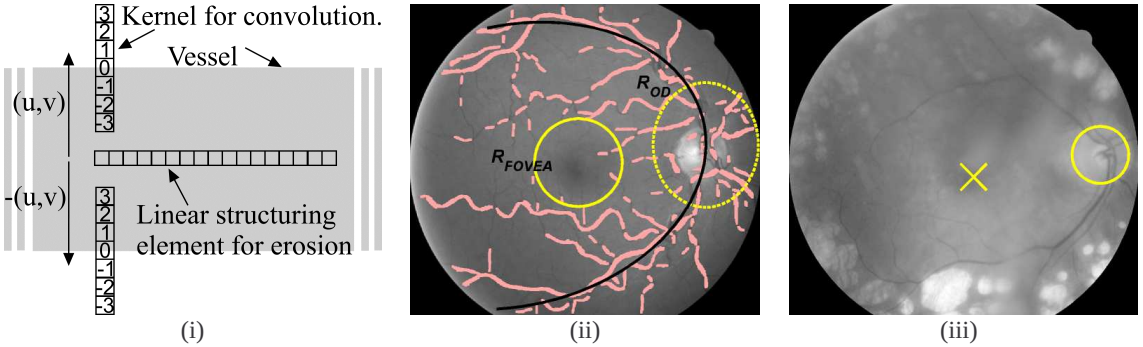


Figure 2. (i) Linear and morphological filters for vessel detection (ii) Results of vessel segmentation (pink), Hough transform applied to major vessels (black semi-ellipse) and approximate localisation of OD and fovea (yellow). (iii) Accurate localisation of OD and fovea (image contains laser photocoagulation treatment scars).

3.3. Optic disc and fovea location

A region, R_{OD}, reliably containing the OD, is a region of height 520 by width 440 pixels centred on the point of the semi-ellipse with vertical tangent, (Figure 2(ii)). The OD location is then determined more accurately using a Hough transform searching for the expected circular shape of the OD applied to edges in the red and green planes of the image. The maximum point in Hough space within R_{OD} is taken as the centre of the OD, (Figure 2(iii)).

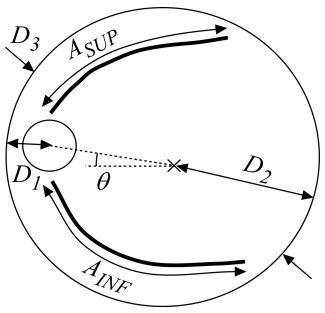
A region R_{FOVEA}, found to reliably contain the fovea, is a circle of diameter 400 pixels centred on a point which lies on a line through the centre of the OD and the centre of the semi-ellipse at a distance of 585 pixels from the OD, (Figure 2(ii)). The fovea location is determined more accurately by band-pass filtering the image, detecting minima and calculating the correlation coefficient with a fovea model at each minimum. The minimum with highest correlation coefficient within R_{FOVEA} is taken as the fovea location, (Figure 2(iii)). If all minima are external to R_{FOVEA} then the centre of R_{FOVEA} is used as the fovea location.

3.4. Small pupil artefact (SPA) detection

A SPA, (Figure 1(vii,viii)), is caused when the "SP" (small pupil) feature of the fundus camera is used. The SPA appears as the outer rim of the image having zero or very low contrast, either very dark or coloured green. After suitable pre-processing to reduce the influence of smaller scale features such as vessels, gradient detection is performed to find image regions where the intensity gradient is approximately radial and above a given threshold. This region is then skeletonised, short branches are removed, and sections are removed whose direction is not tangential to a circle centred on the image. If the remaining skeleton pixels form at least 52% (determined using the training set images) of a circle centred on the image then a SPA is assumed to be present.

3.5. FOV evaluation

Metrics were defined based on the locations of the fovea, OD and image edge, defined as the edge of the background or the edge of the SPA if present, as shown in Figure 3.



Symbol	Description	Requirement
D_1	Distance from centre of OD to edge of image.	>0.5*DD*
D_2	Distance of fovea from edge of image.	>2*DD*
θ	Angle of line joining centre of OD to fovea.	>–5.7°, <24.7°
A_{SUP}	Length of superior arcade.	>2*DD*
A_{INF}	Length of inferior arcade.	>2*DD*
D_3	Diameter of visible fundus.	≥1400 pixels

Figure 3. Metrics used in the determination of adequate FOV. *DD* is the mean OD diameter (246 pixels).

4. Results

The results for the accuracies of feature detection are shown in Table 2. Accuracy of right/left (R/L) eye determination is presented as the percentage of images in which the correct result was obtained according to the clinical grader: all errors were images with no OD visible. Accuracy of location of fovea and OD are presented as the percentage of images in which the automatically determined location and the manually determined location are separated by less than the given distance, where *DD*=mean OD diameter (246 pixels): results are limited to images in which the clinical grader deemed these features to be visible. Arcades incomplete and SPA present are presented as sensitivities and specificities for these conditions being detected by the automated system. The clinical grader deemed that a SPA was present in 21 images of the training set and in 53 of the test set. Results for the determination of adequate FOV are given in Table 3. The clinical grader deemed that 27 images of the training set and 50 images of the test set had an inadequate FOV.

Data set	R/L eye	OD location		Fovea location		Arcades incomplete		SPA present	
	Accuracy	Accuracy (<0.5*DD*)	Accuracy (<0.25*DD*)	Accuracy (<1*DD*)	Accuracy (<0.5*DD*)	Sensitivity	Specificity	Sensitivity	Specificity
Training	98.7%	97.4%	95.4%	99.7%	94.7%	61.5%	97.6%	100%	100%
Test	98.3%	98.0%	97.2%	98.7%	95.7%	63.6%	98.6%	100%	99.4%

Table 2. Feature detection performance.

Data set	Sensitivity	Specificity
Training	88.9% (CI: 71.9% – 96.2%)	95.1% (CI: 92.4% – 96.9%)
Test	92.0% (CI: 81.2% – 96.8%)	92.8% (CI: 91.1% – 94.2%)

Table 3. Sensitivity and specificity for detection of inadequate FOV.

5. Conclusions

In the context of an automated image grading system in a DR screening programme, image quality assessment is important so that images of insufficient quality for automated DR grading are passed to manual graders. This study has shown that it is possible to automatically detect images whose FOV is inadequate to a sensitivity of 92.0% and specificity of 92.8%.

Of the 4 images incorrectly classified as adequate FOV by the automated system, 3 were of inadequate clarity for DR grading and so would be picked up by an automated system for assessing image clarity. It is intended that the work described here be combined with automatic assessment of image clarity, to enable a full assessment of the gradability of the image for retinopathy. This will be covered in a future publication.

Acknowledgements

This work is funded by the Chief Scientist Office of the Scottish Executive Health Department.

References

1. J.H. Hipwell, F. Strachan, J.A. Olson, K.C. McHardy, P.F. Sharp, J.V. Forrester, "Automated detection of microaneurysms in digital red-free photographs: a diabetic retinopathy screening tool", *Diabetic Medicine* **17**, pp588-594, 2000.
2. D. Usher, M. Dumskyj, M. Himaga, T.H. Williamson, S. Nussey, J. Boyce, "Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening", *Diabetic Medicine* **21** pp84-90, 2004.
3. D. Usher, M. Himaga, M.J. Dumskyj, J.F. Boyce, A. Sabate-Cequier, T.H. Williamson, E. Mensah, E.M. Kohner, S.S. Nussey and J. Marshall, "Automated assessment of digital fundus image quality using detected vessel area", In *Proceedings of Medical Image Understanding and Analysis*, pp. 81-84, BMVA, 2003.

Developing a Method of Automatic Prostate MRI Segmentation

P. D. Allen,* J. Graham, D. C. Williamson and C. E. Hutchinson.

Imaging Science and Biomedical Engineering, University of Manchester, Oxford Road, M13 9PT, UK.

Abstract.
In patients with Benign Prostatic Hyperplasia (BPH), disease progression can be quantified by measuring the ratio of the volume of prostate's peripheral zone (PZ) to the volume of the remaining central gland CG. Here we report on progress toward automating this measurement by developing a method of automatically segmenting MR images of prostates into PZ and CG using a combination of grey-level tissue classification and 3D shape modelling. A rationale for this approach is presented based on an anlysis of the process of manual segmentation, and its potential limitations are explored for a data set of 22 patients.

1 Introduction

Benign Prostatic Hyperplasia (BPH) is a non cancerous enlargement of the prostate affecting 70% of men between the ages of 61 and 70, rising to 80% for men over 80 [1]. This enlargement can cause constriction of the urethra which runs though the center of the prostate leading to difficulties in passing urine, and can cause the prostate to impinge on the bladder leading to an increased need to urinate. In 20-30% of men aged 80 treatment is required and this can take two forms: surgery to increase the width of the urethra, or drugs to shrink the prostate and/or reduce the smooth muscle tone [2].

Figure 1 shows a typical MR axial slice through the middle of a prostate. Anatomically the prostate is usually divided into several zones, but in an MR image only two regions can be distinguished: the Central Gland (CG), and the Peripheral Zone (PZ) [3]. BPH mostly occurs in the CG region, causing it to expand and compress the PZ region, and this effect becomes more pronounced as the disease progresses - thus the degree of BPH and hence treatment effectiveness, can be quantified by measuring the ratio of PZ to CG tissue [1,4]. Manual segmentation of the two regions from MR slices is extremely time consuming and the goal of this project is to develop an automatic method of CG/PZ segmentation.

2 The data set

The data collected for this project was acquired using a 1.5T Philips Gyroscan ACS MR scanner (software version NT5.3, Power Track 600, synergy body coil). So far we have concentrated on T2 weighted fat suppressed images as these give good CG/PZ contrast and good separation of the prostate from its surrounding tissue. There are 22 patients and for each there are 50 axial slices with a thickness of 2mm and a resolution of 256x256.

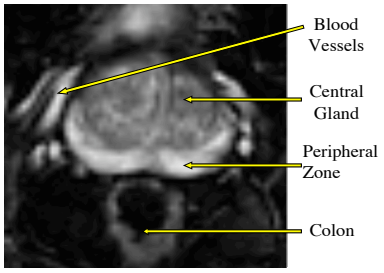


Figure 1. *An axial MR slice of a prostate.*

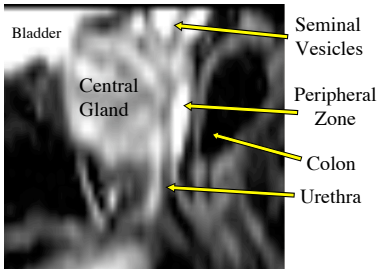


Figure 2. *A sagittal MR slice of a prostate.*

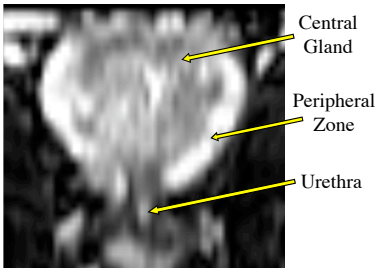


Figure 3. *A coronal MR slice of the prostate.*

Figures 1 to 3 show an example of how the prostate appears in the MR images when sliced in the axial, sagittal, and coronal directions. In T2 weighted images the CG is generally darker than the PZ, and in these images the two can be relatively easily distinguished.

*philip.allen@man.ac.uk

3 Manual segmentation

Manual segmentation of the CG and PZ is usually performed on the axial slices and from the resulting stack of contours a 3D surface can be constructed and hence the CG and PZ volumes can be obtained. Like any manual segmentation this is a tedious and time consuming process and is vulnerable to subjectivity and inaccuracy, however there are also problems peculiar to this data concerning the anatomy of the prostate and the nature of BPH.

3.1 Anatomy

Figure 1 shows an axial slice from the mid-section of the prostate - the view typically published. Here the CG and PZ can be fairly easily distinguished and manual segmentation is reasonably straightforward. Figure 4 shows an axial slice from the same patient but closer to the top of the prostate - in this region the view tends to be complicated by a number of factors: (1) The edges of the prostate are less distinct as the slice is no longer perpendicular to the prostate wall. (2) The bladder is visible and is of similar intensity to the prostate making it difficult to distinguish the two. (3) The seminal vesicles arise from the PZ region and are indistiguishable from the PZ in terms of grey-level alone.

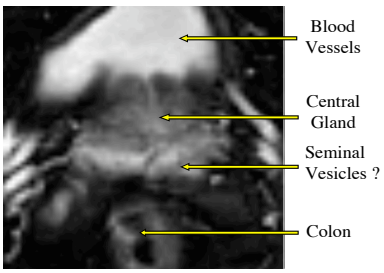


Figure 4. *An axial MR slice of a prostate near the top.*

The most troublesome of these structures are the seminal vesicles as there is no visible boundary between these and the PZ. A difference in texture can be observed in slices further into the seminal vesicles, but in the slices where they first arise from the PZ they are almost impossible to distinguish. This means that the segmenter has to make a decision on where the boundary between PZ and seminal vesicles is based on careful study of the three orthogonal views, and a knowledge of what the prostate should look like - i.e. that it does not have lobes sticking out of it near the top. This is of course vulnerable to considerable subjectivity.

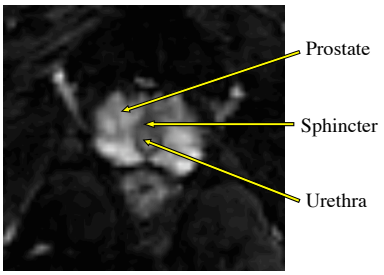


Figure 5. *An axial MR slice of a prostate near the bottom.*

Figure 5 shows another axial slice again from the same patient but toward the bottom of the prostate. The main problem here is that there is a sphincter surrounding the urethra sitting in a recess in the bottom of the prostate (see figure 3) and this can be confused with the CG. Reference to the coronal projection can usually resolve this. A further complication here is that the very tip of the prostate usually splits into two separate lobes and can be confused with other structures below the prostate of similar intensity - once again reference to the other orthogonal projections can help here.

3.2 CG/PZ border ambiguity

As BPH progresses large bright nodules can develop in the CG region giving it a heterogenous appearance and at the same time the pressure exerted on the PZ region can cause it to darken as water is squeezed out of the tissue. This makes the border between the two regions less distinct and in some places non-existent - this is illustrated in figure 6.

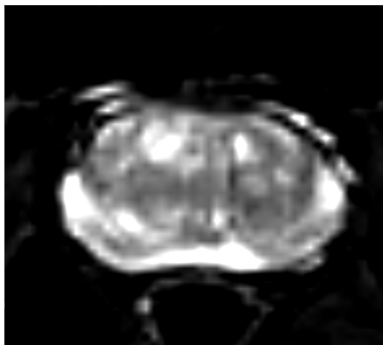


Figure 6. *An axial MR slice of a prostate.*

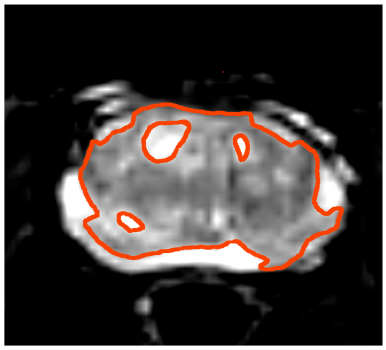


Figure 7. *Manual segmentation of figure 6 based only on the assumption that PZ is brighter than CG.*

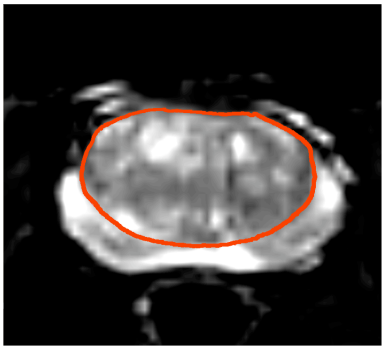


Figure 8. *The clinically correct manual segmentation of figure 6.*

Figure 7 shows the results of segmenting the slice in figure 6 using only the assumption that CG is darker than PZ. Figure 8 shows the same slice but segmented correctly - here the user is using clinical knowledge that the dark regions of PZ are still PZ and that bits of PZ cannot end up within the CG. Thus part of the manual boundary is based on a visible edge between light and dark pixels, and part is based purely on a mental interpolation with no underlying evidence in the image data to support it.

We conclude from this that segmentation is in fact a two stage process: The first stage is to make a coarse classification based on grey level alone and the second is to impose a smooth spatial constraint based on the user's knowledge of what a prostate ought to look like. It is this two-stage interpretation of the segmentation process that we are attempting to formalise in an automatic segmentation method.

4 Automatic segmentation

Our approach is to use grey level pixel classification to mimic the initial coarse segmentation described in section 3.2 and then use a 3D point distribution model to apply a spatial constraint.

4.1 Pixel classification

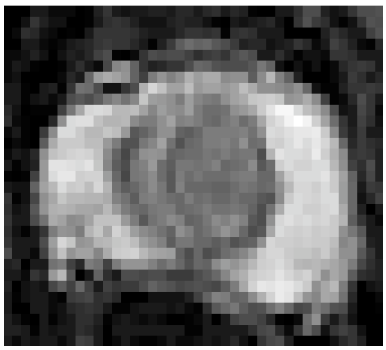


Figure 9. *An axial MR slice of a prostate.*

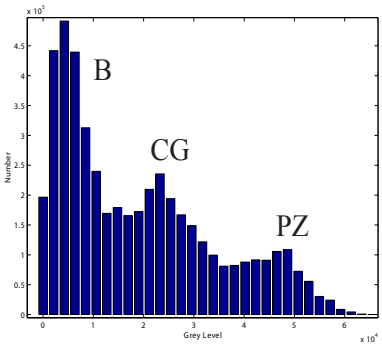


Figure 10. *The grey-level histogram of figure 9.*

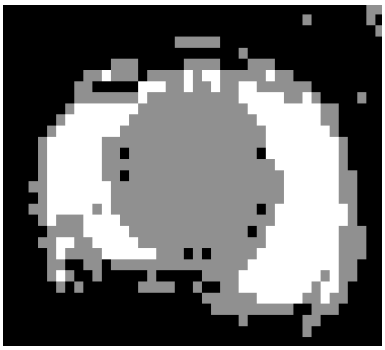


Figure 11. *The result of applying grey-level tissue classification using a three gaussian model to figure 9.*

Figure 9 shows a T2FS axial image cropped around the prostate, and figure 10 shows the grey level histogram for figure 9 in which three distinct peaks can be seen corresponding to surrounding non-prostate tissue which we can now consider as background 'B', central gland 'CG', and peripheral zone 'PZ'. It is a reasonable assumption in an MR image that a pure tissue will produce a mean signal intensity with gaussian distributed noise, and so if we assume that there are only three pure tissues in this image we would anticipate an overall grey level distribution that is a sum of three gaussians.

Figure 11 shows the result of fitting a three gaussian model to the histogram in figure 10 using simplex, and then using the resulting fit to classsify each pixel in the image as either PZ, CG, or B [5, 6]. In this case it works very well as this example has a relatively homogenous PZ and CG and the two are easily distinguished. In a less homogeneous case like the one in figure 6 the bright BPH nodules in the CG will be wrongly classified as PZ and the dark compressed regions in the PZ will be wrongly classified as CG, and so a further smooth spatial constraint is required.

4.2 Shape modelling

For each annotated example of the prostate we have a set of points and from these we can create a 3D point distribution model (PDM) [7]. For this to work effectively the points on each example shape must correspond anatomically and this is usually achieved by using the same number of points when marking up each example and ensuring that the points correspond. However, this is very difficult to do on 3D surfaces and so here we employ a method of automatic correspondence optimisation [8]. In this method the set of surfaces are initially parameterised with a set of equally spaced points spread over the surface. The position of these points are then adjusted until the most compact PDM is achieved - compactness in this case being measured by the 'minimum description length' of the model.

The 3D PDM gives us a method of generating new examples of prostate surfaces using a number of parameters to control shape that is greatly reduced from the number of points involved. This malleable surface then needs to be fitted to the tissue classified data and to do this we need an objective function which measures how well an example shape fits the data. An intuitive approach is to maximmise the number of pixels within each surface to which they have been classified, in other words count the number of pixels classified as CG within the inner surface of the model, the number of PZ pixels between the inner and outer surfaces of the model, and the number of background pixels outside both model surfaces.

5 Preliminary results and discussion

Before attempting to fit the model to the data it is possible to test the theoretical limit of its performance by quantifying its ability to represent each of the shapes in the training set in a leave-one-out test: For each example in the training set a 3D PDM is built from the other examples and the nearest plausible shape to the current example is generated from the model. By repeating this for all examples in the data set and measuring the difference between the real shape and the model generated shape we can quantify the performance of the model. We used the CZ/PZ ratio error as our benchmark as it is what we intend to measure using this method.

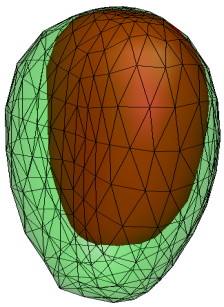


Figure 12. *The mean shape of the double surface 3D PDM of the prostate.*

Most of the examples (19/22) are within 10 $\%$ ($\mu = 7\%, \sigma = 7\%$) of the marked-up shape, but there are a some examples poorly explained by the model giving an error as large as 30%. This limitation appears to be caused by the model being unable to account for extreme variations in the relationship beteween the CG and PZ surfaces. When the two surfaces are modelled separately the error falls to a mean of 2.12% ($\sigma = 1.74\%$) with a maximum of 8%. This suggests that there are too few examples to adequately model both the variation in shape of the two surfaces and their relationship to each other. Modelling the two surfaces individually offers the greatest potential accuracy for this data set, but may reduce the specificity in the model.

Initial experiments in shape fitting using simplex show promise as the objective function landscape in shape space is fairly smooth. However the tendency of simplex to become trapped in local minima requires the search strategy to be specified carefully to achieve optimal results. The introduction of a stochastic element to the search strategy has proven particularly significant and so the use of genetic algorithms will be investigated in the next phase of this project.

Acknowledgements

The 3D PDM code used in this project was written by Tomos Williams. This project is funded by Pfizer.

References

1. A. M. Alam. “Comparison of MR Imaging and Urodynamic Findings in Benign Prostatic Hyperplasia.” *Radiation Medicine* **18**, pp. 123–128, 2000.
2. A. Chauhan. *The use of Active Shape Models for Segmentation of the Prostate Gland from Magnetic Resonance Images*. Master’s thesis, The University of Manchester, Oxford Road, Manchester, UK, M13 9PT, 2001.
3. A. Maio. “Magnetic Resonance Imaging of Prostate Cancer: Update.” *Topics in Magnetic Resonance Imaging* **7(1)**, pp. 54–68, 1995.
4. J. Ishida. “Bening Prostatic Hyperplasia: Value of MR Imaging for Determining Histological Type.” *Radiology* **190**, pp. 329–331, 1994.
5. D. C. Williamson, N. A. Thacker, S. R. Williams et al. “Partial volume tissue segmentation using grey-level gradient.” In *Medical Image Understanding and Analysis*. 2002.
6. K. W. Fleischer, D. H. Laidlaw & A. H. Barr. “Partial-Volume Bayesian classification of material mixtures in MR volume data using voxel histograms.” *IEEE Transactions on Medical Imaging* **17(1)**, pp. 74–86, 1998.
7. T. F. Cootes & C. J. Taylor. “Statistical Models of Appearance for Computer Vision.” www.isbe.man.ac.uk/∼bim/Papers/asm_overview.ps, 2000.
8. R. H. Davies, C. J. Twining, P. D. Allen et al. “Shape Discrimination in the Hippocampus Using an MDL Model.” In *Information Processing in Medical Imaging*, pp. 38–50. Springer-Verlag Heidelberg, 2003.

Detection and Characterisation of the Optic Disk in Glaucoma and Diabetic Retinopathy

R. A. Abdel-Ghafar[a], T. Morris[a*], T. Ritchings[b], I. Wood[c]
[a]Department of Computation, UMIST, [b]School of Computing, Science and Engineering, University of Salford,
[c]Department of Optometry and Neuroscience, UMIST.

Abstract
The optic nerve head is that region of the retina where nerve fibres and blood vessels pass through the sclera. It is sensitive to changes associated with glaucoma that can occur without other symptoms. Diabetes can also be the cause of changes in the appearance of the retina. The changes from the normal appearance can be diagnostic of the diseases, tracking of the changes in sequential images can be used to assess treatment or the progress of the illness. We are concerned with developing automated techniques of generating quantitative descriptions of the retina that might be used in diagnosis and treatment.
Normal and abnormal images were collected from a range of sources, to simulate the mass screening process. Although full colour images were collected, only the green channel was processed as it had the greatest contrast. Blood vessels were removed using morphological techniques. The boundary of the optic disk was estimated using a simple edge detector and the circular Hough transform. Heuristic methods were defined to identify true and false positives. Although we do not expect the optic disk to be circular, this technique identifies it with sufficient accuracy to identify its approximate centre, which is used as the origin of 24 uniformly distributed radial vectors. The images were resampled along these vectors. The resulting image was filtered using the Lee filter and thresholded. In one or other of these images, normal subjects had a region of maximum grey value that extended the full width of the image (at the centre of the optic disk) but did not extend the full height.
Using a chi-squared test, the separation of normal and abnormal images using this test was found to be highly significant ($p < 0.05$, $n = 60$).

1 Introduction
The optic nerve head is that area of the retina where nerve fibres and blood vessels pass through the sclera. It is sensitive to changes in intraocular pressure associated with glaucoma that may occur asymptomatically and which can be diagnostic and that must be tracked to monitor the progress of treatment. Likewise, there are changes in the appearance of the retinal surface that are associated with diabetic retinopathy that are also diagnostic and should also be tracked to allow the progress of the disease and treatment to be assessed. Currently, a clinician, or clinicians, performs the diagnosis and uniform standards cannot be guaranteed. Likewise, a clinician or clinicians perform the tracking at repeat examinations and the repeatability of the assessments cannot be guaranteed. We are seeking methods of quantifying the optic disk objectively: firstly to identify groups of patients who have abnormal ocular appearances and should be further investigated; and secondly to track changes in the appearance of the retina.
In this paper we shall provide a brief overview of previous attempts at quantifying the optic nerve head, we shall describe the data capture process that simulates a mass screening process, we shall describe the algorithms used to analyse the images and present sample results.

2 Review
One of the symptoms of glaucoma is an increase in pressure within the eye as a result of blockage of the flow of aqueous humour, a watery fluid produced by the ciliary body. The increase in pressure damages the optic nerve that carries information from the retina to the brain. In most cases the damage occurs asymptomatically, i.e. before the patient notices any changes to his or her vision. The damage is irreversible; treatment can only reduce or prevent further damage. Visually, the damage is observed as a change in the relative areas of the optic disk and the cup within the disk, figure 1.
Diabetes mellitus is a disorder of carbohydrate metabolism characterised by dysfunctional production of insulin, and thereby elevated blood sugar levels. This leads to problems with different organs such as the eyes, kidneys, heart and cardiovascular system and nerves. Diabetes mellitus has two side effects on the blood vessels: leaking and clogging that can lead to degradation of the retina (retinopathy). This is observed as the proliferation of new vasculature and clotting, figure 2. Retinopathy is one of the main causes of blindness in the working age population [1].

* Currently: Department of Computation, UMIST, PO Box 88, Manchester, M60 1QD. To become School of Informatics, Manchester University, on October 1st 2004.

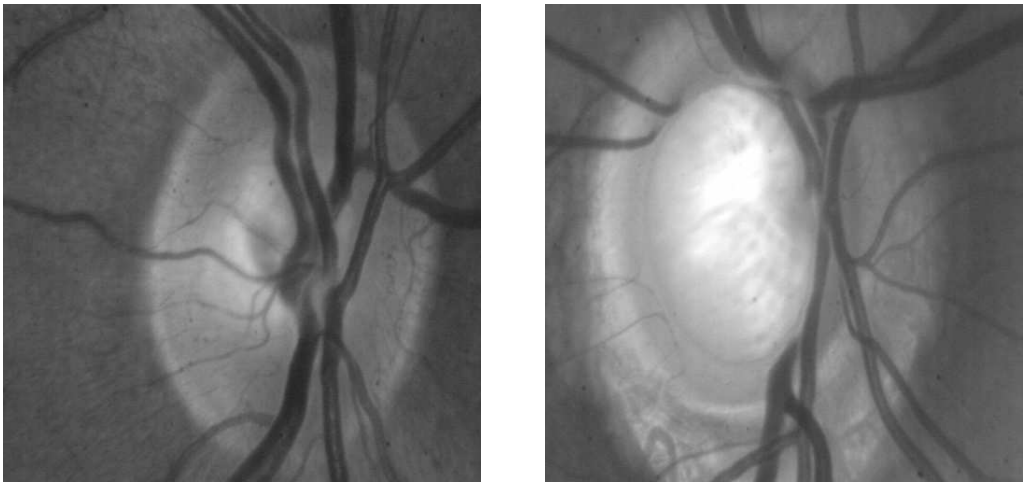


Figure 1. Normal and glaucomatous optic disks.

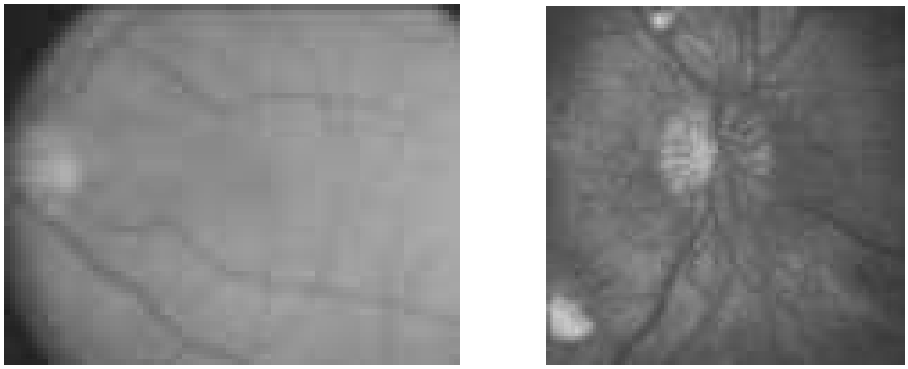


Figure 2. Retina in diabetes mellitus.

To characterise the optic disk (and retina) the optic disk and retinal vasculature must be identified. The disk has previously been identified using the maximum variance as an indication for the location of the disk [2]; template matching followed by principle component analysis (PCA) [3]; by measuring the strength of vessels and the attributes of the bright areas of the image [4] or by identifying the origin of the vessel tree [5]. Colour morphology and dynamic contours (snake) in different colour spaces [6, 7] and wavelet segmentation [8] have been used. Edge detection followed by curve fitting has also been used [9, 10, 10].

3 Materials

The tools to be developed will be used in multiple clinics that are unlikely to have the same image capture devices. To simulate this, our data was gathered from a range of sources. Diabetic retinopathy images were obtained at the Department of Optometry at UMIST, using a Topcon NW6S Non-Mydriatic Retinal Camera. These images were saved as 24-bit true colour JPEG files. Images were taken with a field view of 45 degrees. Glaucoma images were collected from Manchester Royal Eye Hospital; these images were also in the JPEG format. Normal images and a second set of diabetic retinopathy images were downloaded from the STARE (STructured Analysis of the Retina) website [11].

All images were converted to a similar size, as close as possible to 512 by 512 pixels: given the constraints of scaling by an integral factor and retaining the images' aspect ratios. Pixels were represented as 24 bit values. The scaled images were stored in the JPEG format using the best quality settings i.e. near lossless.

Although approximately 90 images were gathered, 16 normal, 31 glaucoma and 13 diabetic retinopathy images were suitable for processing. Of the other 30, some were blurred and others did not contain the whole optic disk. Although full colour images were captured, only the green channel was processed as it was found to have the greatest contrast for these studies.

4 Methods

The green band of the images was processed as it was found that these images had the greatest contrast between the optic disk and the retinal tissue. Firstly, the blood vessels in the image were suppressed by morphological

methods (closing). We then defined 24 radial vectors using the approximate centre of the optic disk as the origin. The image was resampled along these vectors to form a representation that was subsequently processed.

4.1 Optic disk location and image resampling

An approximate location of the centre of the optic disk is required. The image was firstly enhanced using the Sobel operator, and then thresholded using the local mean and variance to compute the threshold value. The remaining points were input to a circular Hough Transform, the largest circle was found consistently to correspond to the optic disk, figure 3. The origin of the circle was used for the resampling vectors.

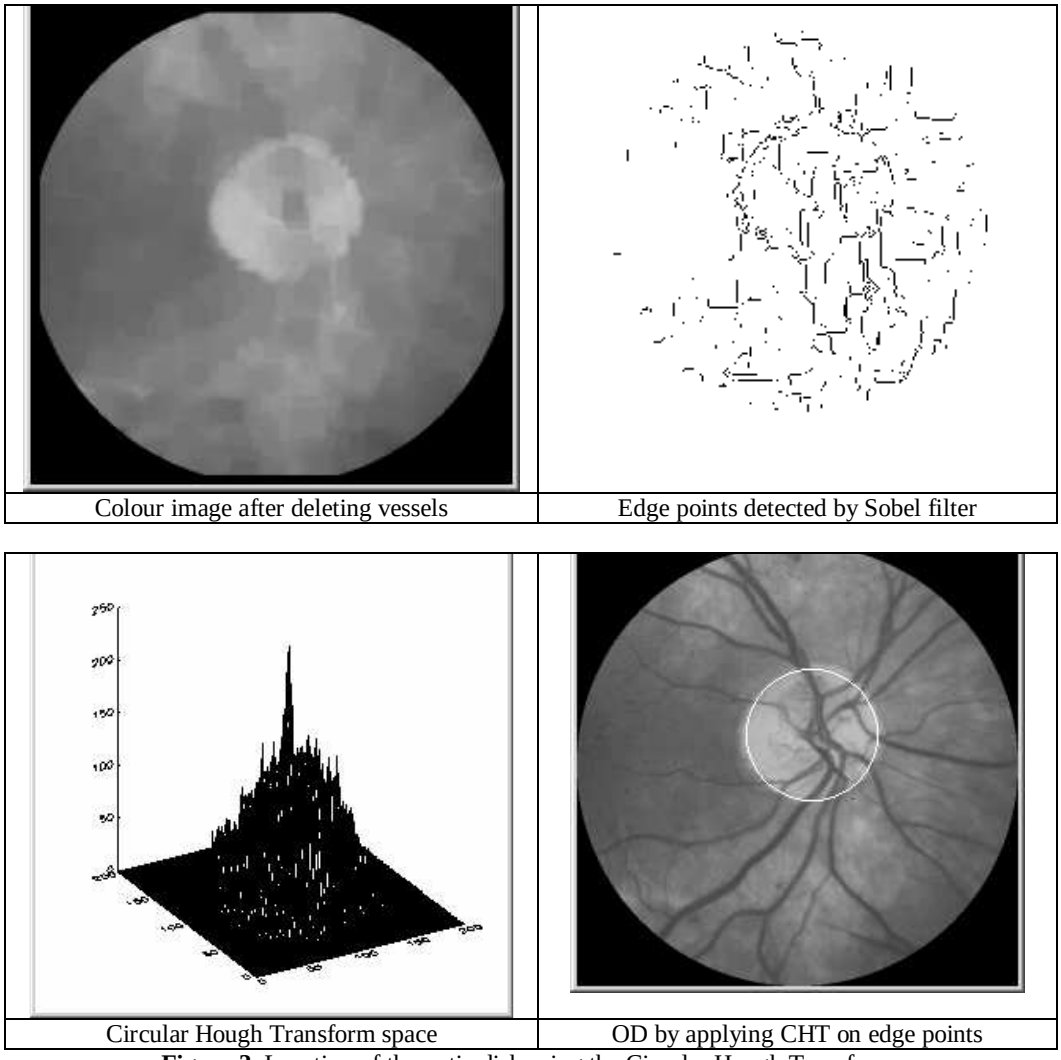


Figure 3. Location of the optic disk using the Circular Hough Transform.

Twenty four uniformly distributed vectors were defined, starting at the optic disk centre. The image was resampled at regular intervals along these using nearest neighbour interpolation.

4.2 Processing resampled images

The resampled images were processed using the Lee filter [12] which suppresses small scale variations whilst retaining significant features. An 11 by 11 pixel kernel was used.

It was found that in many images of normal retinas (set A), the cup region was transformed into a region of maximum intensity spanning the width of the image, but not extending the image's height. This image was also thresholded, at a value of mean + 0.5*standard deviation. It was found that a set of normal images (set B) also had a region of maximum intensity spanning the width of the image, but not extending the image's height. The union of the two sets includes virtually all of the normal images.

5 Results

Sixty images, including 15 normal and 45 abnormals were processed as described above. If the resultant images were members of set A or set B they were classified as normals. The results are summarised in the contingency table, table 1.

		Hypothesis	
		Normal	Abnormal
Result	Positive	9	6
	Negative	7	38

Table 1. Contingency table.

A chi-squared test gave a significant result ($p < 0.05$), indicating that the test is able to separate normal from abnormal (glaucoma or diabetic retinopathy) images. A success rate of 65% was obtained, with sensitivity and specificity rates of 60% and 84% respectively.

6 Conclusions

In this study we set out to develop methods of separating normal from abnormal images (cases of glaucoma or diabetic retinopathy). These would be used in a screening clinic to identify at-risk patients.

Images were collected from various sources to mimic the data collected at a range of sites. Methods were developed to separate the normal from the abnormal images; this was done with reasonable success. Whilst the modest success could be attributed to the insensitivity of our analysis, it can also be attributed to the nature of the diagnosis: we are labelling images as being abnormal or not, without recognising that there is a spectrum of appearances.

The tests indicate that the optic disc's appearance is more uniform in the normals and becomes progressively less so as the diseases progress.

Future work is directed in two directions: accumulating further data and developing more robust and accurate methods of processing this highly variable data.

References

[1] I. Constable, M. McCombe, P. Mitchell, J. O'Day, P. Phillips, A. Stocks, H. Taylor, and T. Welborn, *Diabetes and the Eye - Professional Guidelines*: Australian Diabetes Society for Diabetes Australia, 1996.

[2] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson, "Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images", *British Journal of Ophthalmology*, vol. 83, pp. 902-910, 1999.

[3] H. Li and O. Chutatape, "Automatic Location of Optic Disk in Retinal Images", presented at Image processing, Thessaloniki, Greece, pp. 837-840, Oct, 2001.

[4] S. Chaudhuri, N. K. Chatterjee, and M. Goldbaum, "Automatic Detection of the Optic Nerve in Retinal Images", presented at Proceedings IEEE International Conference on Image Processing, Singapore, pp. 1-5, 1989.

[5] A. Hoover and M. Goldbaum, "Locating the Optic Nerve in a Retinal Image Using the Fuzzy Convergence of the Blood Vessels", *IEEE Transactions on Medical Imaging*, accepted October 2002, to appear.

[6] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Colour morphology and snakes for optical disk localisation", presented at Medical Image Understanding and Analysis, Portsmouth, pp. 21-24, Jul, 2002.

[7] F. Mendels, C. Heneghan, and J. Thiran, "Identification of Optic Disk Boundary in Retinal Images Using Active Contours", presented at Irish Machine Vision and Image Processing Conference, pp. 103-115, 1999.

[8] T. Morris and Z. Newell, "Segmentation of Retinal Images Guided by the Wavelet Transform", presented at Image and Signal Processing, Manchester, pp. 303-306, Nov, 1996.

[9] G. Zahlmann, B. Kochner, I. Ugi, D. Schuhmann, B. Liesenfeld, A. Wegner, M. Obermaier, and M. Mertz, "Hybrid Fuzzy Image Processing for Situation Assessment", *IEEE Engineering in Medicine and Biology Magazine*, vol. 19, pp. 76-83, 2000.

[10] K. G. Goh, W. Hsu, M. Li Lee, and H. Wang, "ADRIS: An Automatic Diabetic Retinal Image Screening System", in *Medical Data Mining and Knowledge Discovery*, vol. 60, *Studies in Fuzziness and Soft Computing*, K. J. Cios, Ed.: Physica-Verlag, pp. 181-210, 2001.

[11] A. Hoover, "STARE project website", retrieved May, 2003, http://www.ces.clemson.edu/~ahoover/stare.

[12] J. S. Lee, "Speckle Suppression and Analysis for Synthetic Aperture Radar Images", *Optical Engineering*, vol. 25, pp. 636-645, 1986.

KCCA for fMRI Analysis

David R. Hardoon[a*], John Shawe-Taylor[a] and Ola Friman[b]

[a]Southampton University, Southampton, UK., [b]Harvard Medical School, Boston, USA[†]

Abstract. We use Kernel Canonical Correlation Analysis (KCCA) to infer brain activity in functional MRI by learning a semantic representation of fMRI brain scans and their associated activity signal. The semantic space provides a common representation and enables a comparison between the fMRI and the activity signal. We compare the approach against Canonical Correlation Analysis (CCA) by localising "activity" on a simulated null data set. Finally we present an approach to reconstruct an activity signal from a testing-set fMRI scans (both simulated and real), a method which allows us to validate our initial analysis.

1 Introduction

Understanding the functional processes of the brain is still a new and difficult task. Functional Magnetic Resonance Imaging (fMRI) is a relatively new tool with the purpose of mapping the sensor, motor and cognitive tasks to specific regions in the brain. The underlying mechanics of this technique is in the regulation of the blood flow as an excess of oxygen is supplied to active neurones causing an increase in oxygenated blood surrounding the tissue of the active brain region. This effected is referred to as BOLD (Blood Oxygenation Level Dependent) signal.

We present a Kernel CCA (KCCA) approach to measure the active regions of the brain using fMRI scans and their activity signal. Friman et. al [1] have shown that CCA can gives us the ability to introduce several time-courses as the BOLD response has been shown to vary both between people and brain regions. In previous work [2] we have shown that applying kernel methods [3] can increase the performance of CCA. Finally we show that due to the properties of KCCA [4] we are able to use this approach to reconstruct the activity signal from an "unknown" testing-set fMRI scans a process that allows us to validated our prior analysis.

The paper is divided as follows, Section 2 gives a brief introduction the method used and the baseline comparison. Section 3 describes the experiments taken place and Section 4 brings forward final discussion.

2 Method

Proposed by H. Hotelling in 1936 [5], Canonical correlation analysis seeks a pair of linear transformations one for each of the sets of variables such that when the set of variables are transformed the corresponding co-ordinates are maximally correlated [6]. Let

$$\rho = \max_{\mathbf{w}_x, \mathbf{w}_y} \frac{\mathbf{w}_x' C_{\mathbf{xy}} \mathbf{w}_y}{\sqrt{\mathbf{w}_x' C_{\mathbf{xx}} \mathbf{w}_x \mathbf{w}_y' C_{\mathbf{yy}} \mathbf{w}_y}} \tag{1}$$

the maximum canonical correlation is the maximum of ρ with respect to $\mathbf{w}_x$ and $\mathbf{w}_y$. We represent the two time-courses as a linear combination of pixel time-course

$$\mathbf{x}(\mathbf{t})\mathbf{w}_\mathbf{x} = x_1(t)w_{x_m} + \ldots + x_m(t)w_{x_m}$$

and any chosen time sequence to represent the fMRI modal

$$\mathbf{y}(\mathbf{t})\mathbf{w}_\mathbf{y} = y_1(t)w_{y_m} + \ldots + y_m(t)w_{y_m}.$$

CCA may not extract useful descriptors of the data because of its linearity. Kernel CCA offers an alternative solution by first projecting the data into a higher dimensional feature space (where $n < N$)

$$\phi : \mathbf{x} = (\mathbf{x}_1, \ldots \mathbf{x}_n) \mapsto \phi(\mathbf{x}) = (\phi_1(\mathbf{x}), \ldots, \phi_N(\mathbf{x}))$$

before performing CCA in the new feature space, essentially moving from the primal to the dual representation approach.

*Correspondence author (Hardoon) at drh@ecs.soton.ac.uk.

†The work was done while Ola Friman was at Linkoping University, Sweden

Definition 1. $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product of the vectors $\mathbf{x}$,$\mathbf{y}$ and is equal to $\mathbf{x}'\mathbf{y}$. Where we use A' to denote the transpose of a vector or matrix A.

Kernels are methods of implicitly mapping data into a higher dimensional feature space, a method known as the "kernel trick". A kernel is a function K, such that for all $x, z \in \mathcal{X}$

$$K(x, z) = \langle \phi(x), \phi(z) \rangle \tag{2}$$

where ϕ is a mapping from $\mathcal{X}$ to a feature space F. The weights can be written as a linear combination of the training examples, let

$$\mathbf{w}_x = \phi(\mathbf{x})\alpha. \tag{3}$$

Hence we obtain from CCA (for details [4, 6]) the dual representation

$$\rho = \max_{\alpha,\beta} \frac{\alpha' K_x K_y \beta}{\sqrt{\alpha' K_x^2 \alpha \beta' K_y^2 \beta}}. \tag{4}$$

As KCCA requires two views we provide K_x as a kernel from the fMRI brain scan and K_y as a kernel from the chosen time-sequence. In [6] we observe that with full rank kernel matrices maximal correlation can be obtained, suggesting that learning is trivial. To force non-trivial learning we introduce a control on the flexibility of the projection mappings by penalising the norms of the associated weights (detailed description of CCA and KCCA can be found in [4]), we obtain

$$\rho = \max_{\alpha,\beta} \frac{\alpha' K_x K_y \beta}{\sqrt{(\alpha' K_x^2 \alpha + \kappa \alpha' K_x \alpha)(\beta' K_y^2 \beta + \kappa \beta' K_y \beta)}}. \tag{5}$$

3 Experiments

3.1 Activity Localisation

We compare KCCA to the baseline CCA as presented in [1, 7, 8, for brevity we refer the reader to the papers for details on the CCA method]. The regularisation parameter κ from equation (5) is computed a priori as described in [2, 6]. In the following experiment we use the correlation values computed by CCA though in KCCA we prefer to compute the weights associated to the pixels, as this can give us more information on the activity of each pixel, this step can also be done with CCA although we do not compute it as this approach is not as intuitive as with KCCA.

As it is impossible to tell which method is better when real data is used, we experiment with controlled simulated data. We embed square-wave "activity" in a nulldata set (no brain activity). The paradigm of the applied activity is 10 images rest, 10 images activity and so forth. Resulting with 200 time points. As we know the activation period we use for our time sequence a square-wave representation of activity (1) and rest (-1) over the 200 time-course. We use a linear kernel for both the fMRI data and the square-wave "activity".

In figure 1 the found true-positive and false-positive pixels using CCA are plotted we are able to observe that although there is a decline in the number of false-positive pixels located, the number of false-positive outnumbers the number of true-positive pixels. At a threshold of 0.74 we are able to observe in figure 1(right plot) that the rate of false-positive drop below that of the true-positive but not in a significant measure. In figure 2 the found true-positive and false positive pixels using KCCA are plotted. As expected the number of false-positive pixels start at a much higher rate then that found with CCA as with KCCA we take into account all the pixels in the image. Although this worse start we are able to observe a sharp drop in the number of false positive pixels accompanied with a stedy and slower drop in the number of true-positive pixels. We also find that the number of true-positive pixels located surpasses the number of false-positive, with a relatively low threshold. This suggests that KCCA is able to extract a better ratio of true to false positive pixels.

3.2 Statistical Reconstruction

In the following section we present an approach of statistically reconstructing a signal from the fMRI scans. This reconstruction approach will allow us to determine the validity of our prior analysis for if we have learnt the

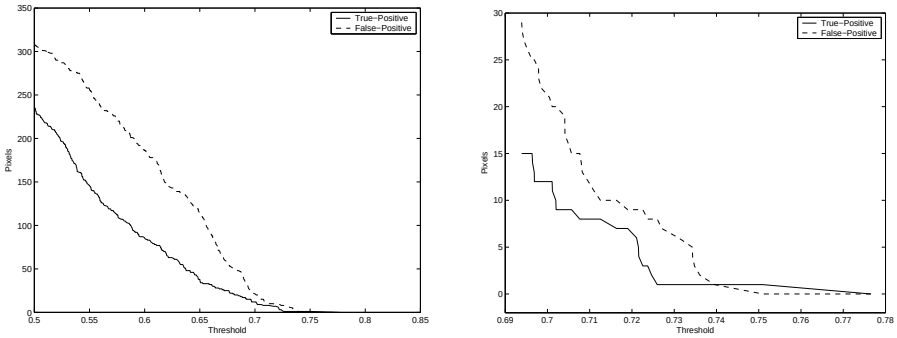


Figure 1. Left: **CCA Found True/False-Positives**. Right: **Same Plot, Zoomed-In.**

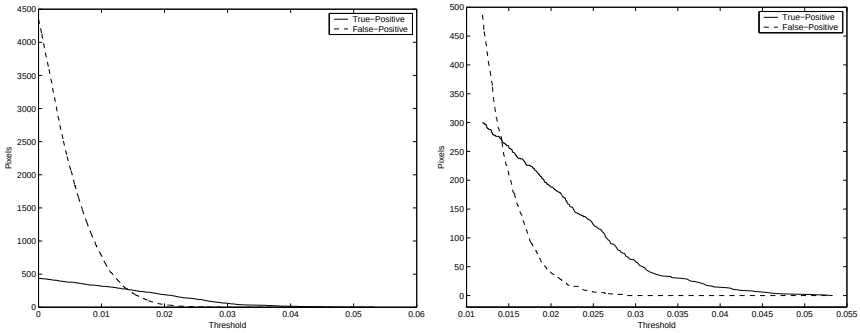


Figure 2. Left: **KCCA Found True/False-Positives**. Right: **Same Plot, Zoomed-In.**

appropriate function we will be able to reconstruct it. Let X be the fMRI training examples and X_t the fMRI testing examples, Y be the training activity time sequence and Y_t the testing activity signal we want to reconstruct. Let

$$g_{w_x,w_y} = \|Xw_x - Yw_y\|^2 \tag{6}$$

where $g_{w_x,w_y} \approx 0$ as we want the feature Xw_x from one view of the data to be identical to the feature Yw_y obtained from the second view of the data, this will be true on the training data if there is a high correlation between the two views. Therefor we can rewrite equation (6) as

$$\|Xw_x - Yw_y\|^2 \approx 0 \tag{7}$$

$$Xw_x \approx Yw_y \tag{8}$$

We divide the fMRI data into a training and testing set. On the training data we compute the KCCA coefficients α, β. Let $K_{x_t} = \langle X_t, X \rangle$ be the fMRI testing kernel and $K_{y_t} = \langle Y_t, Y \rangle$ be the time sequence testing kernel. [4] have shown that this equivalence can be held true also for the testing data using efficient regularisation. Hence we justify the usage of equation (8) and equation (3) to define

$$\begin{aligned} X_t w_x &\approx Y_t w_y \\ K_{x_t}\alpha &\approx Y_t' Y \beta. \end{aligned}$$

As we are interested in finding the testing-set unknown activity time sequence we can rearrange the equation to

$$Y_t \approx (K_{x_t}\alpha \cdot (Y\beta)^{-1})'. \tag{9}$$

As we are no longer interested in the weight vector but in the reconstruction of the signal, we are not confined to the usage of inner product kernels, in the following experiment we compare the success rate between the linear kernel as used, to the Gaussian kernel (defined in equation (10))

$$K(x_i, x_j) = \exp \frac{-\|x_i - x_j\|^2}{\sigma^2} \tag{10}$$

using σ as the minimum distance between the different labelled images.
We test our approach using the square-wave time sequence of 1 representing activity and -1 representing rest for both the simulated and real data experiments. The real data is comprised of mental calculation, the adding of two numbers, and right hand index finger flexing. For the simulated data and mental calculation we use the first 160 scans for training and the remaining 20 for testing, while with the finger flexing we use the first 180 for training and the remaining 20 for testing. We randomise the examples prior to the training and testing separation. Once we obtain the reconstructed Y_t we threshold it by $T = 0$, i.e. Let $\hat{Y}$ be the thresholded reconstructed signal

$$\hat{Y} = \begin{cases} 1 \ if \ Y_t \geq 0 \\ 0 \ otherwise \end{cases}$$

Table 1 shows the average overall results of successfully reconstructing the activity time sequence for the testing fMRI data over 100 repeats using both a linear & Gaussian kernels. We are able to see that the linear kernel performs better then the Gaussian. It is important to stat at this stage the difference between CCA and kernel CCA (KCCA) with a linear kernel. The former uses a larger number of features than in CCA, which are computed implicitly in the kernel.

Table 1. Success rate in reconstructing the test activity time sequence

Data-Set	Linear	Gaussian
Simulated data	97.4%	95.8%
Finger flexing	75.25%	69.25%
Mental calculation	48.6%	41.55%

We provide an initial experiment in attempt to learn the mental process prier to the finger flexing by setting the square-wave sequence such that the three images before the actual finger flexing were considered as active and all the remaining images were considered as inactive. We have trained as before using a linear kernel and attempted to reconstruct this new square-wave sequence over an average of 100 random repeats. We find that we can successfully reconstruct the signal with an success average of 83.3%.

4 Discussion

For future work we would like to try more elaborate time basis functions and to experiment on different data types (emotional, mental and other motor data) and tailored kernels for better extracting the activity/signal. A further interesting avenue would be to observe the performance of applying our KCCA approach to other techniques of brain analysis and also to more complex tasks. We also speculate that KCCA would be able to handle a multiple task fMRI scenario.

Acknowledgements

The first and second authors would like to acknowledge the financial support of EU Project LAVA, No. IST-2001-34405 and thank Craig Saunders for many useful discussions.

References

1. O. Friman, J. Carlsson, P. Lundberg et al. "Detection of neural activity in functional MRI using canonical correlation analysis." *Magnetic Resonance in Medicine* **45(2)**, pp. 323–330, February 2001.
2. D. R. Hardoon & J. Shawe-Taylor. "KCCA for different level precision in content-based image retrieval." In *Proceedings of Third International Workshop on Content-Based Multimedia Indexing*. IRISA, Rennes, France, 2003.
3. N. Cristianini & J. Shawe-Taylor. *An Introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2000.
4. J. Shawe-Taylor & N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
5. H. Hotelling. "Relations between two sets of variates." *Biometrika* **28**, pp. 312–377, 1936.
6. D. R. Hardoon, S. Szedmak & J. Shawe-Taylor. "Canonical correlation analysis; an overview with application to learning methods." Technical Report CSD-TR-03-02, Royal Holloway University of London, 2003.
7. O. Friman, M. Borga, P. Lundberg et al. "A correlation framework for functional MRI data analysis." In *Proceedings of the 12th Scandinavian Conference on Image Analysis*. SCIA, Bergen, Norway, June 2001.
8. O. Friman. *Adaptive Analysis of Functional MRI Data*. Ph.D. thesis, Linkoping Studies in Science and Technology, 2003.

Comparison of methods for visualising changes from orthognathic surgery

Lucy Miller[a], Elizabeth Berry[a*], David O Morris[b,c]
[a]Academic Unit of Medical Physics and [b]Academic Unit of Child Dental Health, University of Leeds, Leeds, UK; [c]Orthodontics Department, Leeds Teaching Hospitals NHS Trust, Leeds, UK

Abstract. Surface laser scanning can provide information about the three dimensional shape of the face. If scans are acquired before and after orthognathic surgery, they can be processed to allow quantification and visualisation of changes. The purpose of this study was to assess how well four different visualisation methods communicated surgical changes to observers. Three groups were studied: 16 females for skeletal III mandibular advancement; 13 males for bimaxillary skeletal III surgery and a control group of 12 (6 female) who did not undergo surgery. Ten observers twice viewed visualisations generated by the four methods and were asked to determine to which of the three groups each belonged. Sensitivity and specificity were determined for each method and subject group. Intra-observer repeatability was assessed using Cohen's kappa. The measured sensitivity for all the visualisation methods and all the subject groups was low, in the range 20-60%. Specificity was higher (81-100%), but many visualisations were classified as "Don't know" by observers (12-39%). Intra-observer repeatability was good (κ=0.61). None of the methods tested was consistently reliable in communicating information on facial changes to the observer; necessary improvements to the processing and display have been identified.

1. Introduction

Orthognathic surgery is performed on patients who suffer from compromised jaw function, mastication problems and speech difficulties; it involves corrective surgery of the upper, lower or both jaws. The teeth and jaws are manipulated in three-dimensions, to obtain the most aesthetic, stable and functional result. If surface laser scans are acquired before and after surgery, following spatial registration of the surfaces it is possible to quantify the changes that have taken place.[1] This can be useful for assessing soft tissue changes in individual cases [2,3,4,5] and there is also the potential for visualising the average effect calculated from a group that have undergone the same procedure.[6] Several methods have been developed to assist the visualisation of these changes, by calculation of a displacement vector (V) between spatially registered pre- and post-operative scans:

- Correspondence by sensitivity to movement (CSM)[7]: V indicates the direction and distance between the points on the two scans with the most similar curvedness, shape and relative angle.
- Normals: V indicates distance along a perpendicular from points on the post-operative surface to its intersection with the pre-operative surface.
- Radial[8]: V indicates the distance between the intersections with the two surfaces of lines drawn from the centroid of the post-operative surface.
- Closest point: V indicates the distance and direction of the closest point on the pre-operative scan from each point on the post-operative scan.

The magnitude of the vector is then displayed, using a colour scale, on the post-operative surface; an alternative visualisation scheme (not used in this study) is to display the vectors on the post-operative surface using needle-shaped symbols indicating the vector magnitude and direction. Clearly the vectors could also be calculated from the pre- to post-operative surface. In future work it is intended that colour displays of this type will be used by orthognathic surgeons in assessing soft tissue changes in individuals and in predictions using averaged data. However, no published study was found that assessed the reliability of such displays in communicating information to the observer, though a misleading display would not be suitable for clinical use. The aim of this study then was to determine the performance of the four methods of visualisation in communicating surgical changes to observers.

2. Method

Two surgical groups and one control group were studied. Individuals in the surgical groups were under the care of a single Consultant Orthodontist (DOM) in the Leeds Teaching Hospitals NHS Trust. Ethical approval for the study was awarded by the Leeds (West) Research Ethics Committee. Group A comprised 16 females undergoing skeletal II mandibular advancement. Group B comprised 13 males undergoing bimaxillary skeletal III surgery. The control group was made up of 6 female and 6 male non-growing adults who did not undergo orthognathic surgery. Facial laser scanning was performed using a Cyberware 3030HRC laser scanner (Cyberware Inc., Monterey, CA, USA). Each surgical subject had two facial laser scans, the first on the day before surgery and the second four months post-operatively. The control subjects had two scans, acquired two weeks apart (for convenience these will also be referred to as pre- and post-operative scans).

* Corresponding author: e.berry@leeds.ac.uk

The acquired data were converted to a triangular surface mesh in the Stanford PLY format,[9] and the number of points then reduced to 15% of the original number which resulted in a surface with the same geometry but very much reduced in size to facilitate the later processing.[7] The forehead region of the face does not change as a result of routine orthognathic surgery or natural growth beyond nine years of age,[2,8,4] so spatial registration of the pre- and post-operative scans was performed by aligning the forehead segmented from the pre-operative scan with the face acquired post-operatively. A number of different registration algorithms were available,[10] and just one of these was used for this study to reduce the number of variables. The forehead and orbital region of the face from each pre-operative scan was aligned with the corresponding post-operative scan using the iterative closest point algorithm (ICP).[11] The resulting transform was applied to the whole face of the pre-operative scan. The differences between registered pairs of scans were then analysed using the four techniques outlined previously. The magnitude of V at each point was indicated on the post-operative scan using a colour scale (Figure 1). Warm colours (yellow, orange, red) represented backwards movement, and cold colours (green, blue, purple) represented forwards movement. Areas with no change appeared grey.

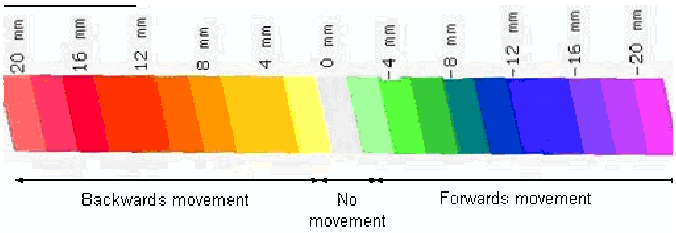


Figure 1. The colour millimetre scale associated with the visualisations

The four visualisations for each subject were randomly ordered and shown to 10 orthodontists in training. The observers were blind to the study group, but received instruction to understand what each type of visualisation represented. Each observer decided which operation, if any, the subject had undergone. They could also give a "Don't know" response. The observers repeated their observations on the same images but in a different random order, one to two weeks later. Sensitivity and specificity for identification of the surgical procedure, with 95% confidence limits, were calculated from their responses. The proportions of "Don't know" responses for each subject group and for each visualisation method were determined to allow assessment of how many of the visualisations could not be interpreted at all. Intra-observer agreement was assessed using Cohen's kappa, adjusting for the agreement that could occur by chance i.e. Cohen's kappa is the chance-corrected proportional agreement.[12]

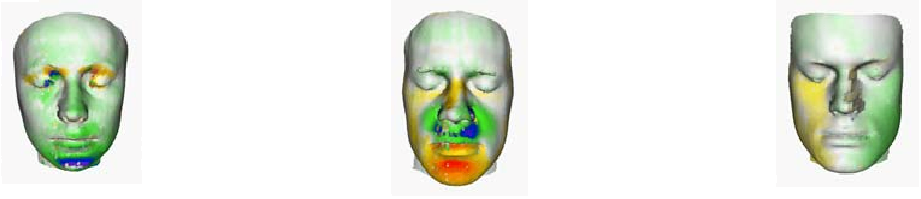
3. Results

The registration process failed on three subjects and they were removed from the study. This reduced the size of Group A to 13. The sensitivity and specificity results are shown in Table 1.

	Sensitivity/Specificity (%) 95% confidence intervals			
	CSM	Normals	Radial	Closest point
Group A	58/98 49-66 / 98 -100	57/97 48-65 / 95-99	59/92 46-71 / 89-96	51/96 42-59 / 94-98
Group B	22/100 15-29 / 100-100	26/100 19-34 / 99-100	24/100 17-31 / 100-100	20/100 13-27 / 100-100
Group C	56/86 47-65 / 82-90	43/92 34-51 / 89-96	28/96 20-36 / 93-98	60/81 51-69 / 76-86

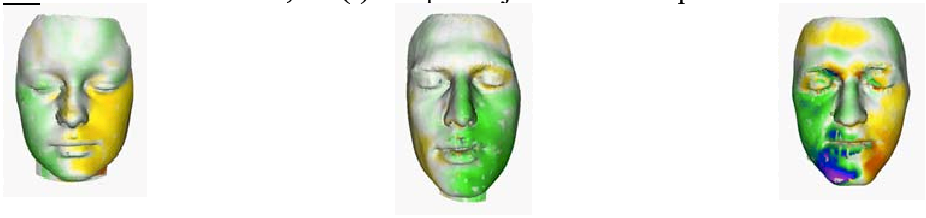
Table 1. Sensitivity and specificity for the four visualisation methods

No method performed best for all the procedures, and although highly specific, the highest sensitivity was only 60%. In Figure 2 an example from each subject group is shown, the example is one that was well interpreted by the observers using the visualisation method that had the highest sensitivity for that group. The low sensitivities occurred because many images were poorly interpreted, and in Figure 3 a poorly interpreted example from each study group is shown, in each case the visualisation method is the one with the lowest sensitivity for that group. A summary of the "Don't know" responses is shown in Table 2. Kappa values are shown in Table 3. These results indicate good intra-observer agreement (The strength of agreement is defined as "Good" for values of kappa in the range 0.61-0.80.[12]



(a) (b) (c)

Figure 2. Examples of well interpreted visualisations. (a) Group A subject - radial method, (b) Group B subject - normals method, and (c) Group C subject - the closest point method.



(a) (b) (c)

Figure 3. Examples of poorly interpreted visualisations. (a) Group A subject - closest point method, (b) Group B subject - closest point method, and (c) Group C subject - radial method.

	"Don't know" response / %			
	CSM	Normals	Radial	Closest point
Group A	20.8	20.8	23.1	12.3
Group B	38.5	31.5	33.8	36.2
Group C	36.7	32.5	39.2	23.3
Overall	31.8	28.2	31.8	23.9

Table 2. Percentage of responses from observers classified as "don't know".

	Kappa	Standard error	95 % confidence interval
Group A	0.70	0.023	0.65-0.74
Group B	0.70	0.034	0.63-0.77
Group C	0.62	0.024	0.58-0.67
Overall	0.61	0.015	0.58-0.64

Table 3. Kappa values for intra-observer repeatability

4. Discussion

The measured sensitivity for all the visualisation methods and all the subject groups was low, in the range 20-60%, thus none of the visualisation methods was particularly successful in correctly communicating the surgical change that had taken place. The specificities were in each case higher than the sensitivity, and were in some cases 100%, meaning that the visualisation methods were better at indicating what had *not* been done.

The radial visualisation method has been favoured by McCance et al, [2,13,14,8] who used it to visualise the facial changes from orthognathic surgery in Class I, II and III subjects. In our study the radial method had a higher sensitivity for the Group A subjects than for Groups B and C, although it must be noted that the sensitivity was very close to those for the CSM and Normals method within Group A. Group A subjects are from Class II, so our results support the use of the method in the published work. However, we found that all methods, including the radial method, had sensitivity under 26% for Group B (Class III) subjects, and that the sensitivity for the normals method was marginally higher than that for the radial method. These subjects had undergone bimaxillary surgery, which involves a more complex two-jaw procedure with less predictable soft tissue changes than are seen for mandibular advancement cases.

None of the four methods had the highest sensitivity for all three subject groups. This has implications for the future use of the methods, as it may be necessary to tailor the visualisation method to the type of procedure performed to ensure the best interpretation. However, there are several confounding factors, discussed below, that may impact on this finding. It was established that repeatability of interpretation was good, thus the observers made the same decision on the two occasions they viewed a particular visualisation. The variations in interpretation may be attributed to three factors: errors in registration of the pre- and post-operative scans, a misunderstanding by observers of the meaning of the colour scale, and observer fatigue.

Although subjects were only included in the study if the alignment of the foreheads was judged subjectively to be good, it was noted that in some subjects there appeared to be a residual rotation between pre- and post-operative scans in the lower part of the face. In the visualisations this appears as one side of the face

being warm coloured and the other cold coloured (Figure 3a), which may lead the observer to classify the visualisation as a "Don't know". Errors in registration have arisen from the lack of strong shape characteristics in the segmented foreheads, especially as we chose to use the basic ICP algorithm for this study and not one that incorporates information on, for example, curvatures and normals.[10] In future work the effect of the choice of registration algorithm, and of constraining the match using manually,[15,2] semi-automatically[16] or automatically identified anatomical landmarks will be investigated.

It is suspected that observers may have over-interpreted the colour scale, thus thinking that an operation has been performed on the control subjects. The colour scale indicates movements of between 1 and 3 mm in pale shades of green and yellow. However, a difference between registered scans of this magnitude may have arisen from the inherent limits of the sequence of acquisitions and processing that is performed. A separate study is underway to determine the minimum difference between scans that can be attributed to change rather than to experimental variation. In future work, the colour scale will be adapted so that changes below that limit are not coloured.

This study involved each observer interpreting 152 images on two separate occasions, and as they progressed through this task they may well have become fatigued and made errors or been inconsistent. Future studies will use smaller subsets of images.

We conclude that although colour visualisations appear to be an objective way to present results, there is a strong degree of subjectivity involved in their interpretation. None of the methods tested was consistently reliable in communicating information on facial changes to the observer.

Acknowledgements

This work was performed as part of the Intercalated BSc Clinical Sciences programme at the University of Leeds, and LM acknowledges the support of her sponsor Dr Craig Saxton MD. Previous support from the European Commission (BRPR-CT97-0378) and the Wellcome Trust (047808/Z/96/Z) is acknowledged. We are grateful to A Tennant, E Hensor and D Bowers for statistical advice, AJ Bulpitt for emergency software modifications and thank the ten trainee orthodontists who acted as observers for their essential contributions.

References

1. E.R.Thomson "The use of morphanalysis in orthognathic surgery", *British Journal of Plastic Surgery* 38, pp. 75-83, 1985.
2. A.M.McCance, J.P.Moss, W.R.Wright et. al. "A three-dimensional soft tissue analysis of 16 skeletal class III patients following bimaxillary surgery.", *British Journal of Oral & Maxillofacial Surgery* 30, pp. 221-232, 1992.
3. A.M.McCance, J.P.Moss, W.R.Fright et. al. "Three-dimensional analysis techniques - Part 1: Three-dimensional soft-tissue analysis of 24 adult cleft palate patients following Le Fort I maxillary advancement: a preliminary report.", *Cleft Palate-Craniofacial Journal* 34, pp. 36-45, 1997.
4. J.P.Moss, A.M.McCance, W.R.Fright et. al. "A three-dimensional soft tissue analysis of fifteen patients with Class II, Division 1 malocclusions after bimaxillary surgery.", *American Journal of Orthodontics & Dentofacial Orthopedics* 105, pp. 430-437, 1994.
5. A.M.Coombes, J.P.Moss, A.D.Linney et. al. "A mathematical method for the comparison of three-dimensional changes in the facial surface.", *European Journal Of Orthodontics* 13, pp. 95-110, 1991.
6. R.Marmulla and H.Niederdellmann "A 3D relaxation mesh for soft-tissue prediction in surgical planning of computer assisted repositioning osteotomies", *CARS 2000: Computer Assisted Radiology and Surgery* 1214, pp. 905-909, 2000.
7. E.Guest, E.Berry, and D.Morris "Novel methods for quantifying soft tissue changes after orthognathic surgery", *International Journal of Oral and Maxillofacial Surgery* 30, pp. 484-489, 2001.
8. A.M.McCance, J.P.Moss, W.R.Fright et. al. "Three-dimensional analysis techniques - Part 2: Laser scanning: a quantitative three-dimensional soft-tissue analysis using a color-coding system.", *Cleft Palate-Craniofacial Journal* 34, pp. 46-51, 1997.
9. "The PLY File Format", http://www.cc.gatech.edu/projects/large_models/ply.html, last accessed 11/8/2004.
10. E.Guest, E.Berry, R.A.Baldock et. al. "Robust point correspondence applied to two and three- dimensional image registration", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, pp. 165-179, 2001.
11. P.J.Besl and N.D.McKay "A method for registration of 3-D shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14, pp. 239-256, 1992.
12. D.G.Altman . *Practical statistics for medical research.* Chapman and Hall, London , 1997.
13. A.M.McCance, J.P.Moss, W.R.Fright et. al. "A three-dimensional analysis of bone and soft tissue to bone ratio of movements in 17 Skeletal II patients following orthognathic surgery.", *European Journal Of Orthodontics* 15, pp. 97-106, 1993.
14. A.M.McCance, J.P.Moss, W.R.Fright et. al. "Three-dimensional analysis techniques - Part 3: Color-coded system for three-dimensional measurement of bone and ratio of soft tissue to bone: the analysis.", *Cleft Palate-Craniofacial Journal* 34, pp. 52-57, 1997.
15. T.J.Coward, R.M.Watson, and B.J.J.Scott "Laser scanning for the identification of repeatable landmarks of the ears and face", *British Journal of Plastic Surgery* 50, pp. 308-314, 1997.
16. V.F.Ferrario, C.Sforza, C.E.Poggio et. al. "Distance from Symmetry - A 3-Dimensional Evaluation of Facial Asymmetry", *Journal of Oral and Maxillofacial Surgery* 52, pp. 1126-1132, 1994.

Combined Ultrasound Speckle Pattern Similarity Measures

J. Revell[1]*, M. Mirmehdi[1] and D. McNally[2]

[1]Department of Computer Science, University of Bristol, Bristol, BS8 1UB, UK
[2]Institute of Biomechanics, University of Nottingham, Nottingham, N97 2RD, UK

Abstract. We present an enhanced block matching approach to improve displacement accuracy in ultrasound sequences using a combination of matching measures. The first measure uses the normalised cross correlation for regions of strong signal and the second measure CD_2, specifically for regions of speckle determined by the speckle signal to noise ratio. We also show displacement field results for simulated speckle and *in vitro* data.

1 Introduction

With modern ultrasound machines providing realtime sequence digitisation, motion estimation research in this area for noise filtering, tracking and registration has increased. In this paper we investigate a novel practical alternative to elastography using speckle tracking to infer tissue motion. Our contribution includes applying two speckle pattern similarity measures, adapting to regions of varying signal and noise within a multiresolution framework with displacement processing. We focus on synthetic and *in vitro* interframe and trajectory displacement accuracy.

Scatter occurs when small imperfections (scatterers) cause seemingly random reflections and refractions of the sound wave. Scatterers account for a decrease in image quality, causing blurring and decreased intensity at impedance boundaries, while within the medium they create speckle. The statistics of the signal depends on the density of scatterers, with a large number of randomly located scatterers following a Rayleigh distribution (fully developed speckle). These conditions are seldom met, resulting in different statistical speckle models being used.

Using B-mode images 2D tissue motion can be measured by tracking the movement of the speckle produced by the back scattering of the ultrasound itself. To date, the most popular approaches to speckle tracking use 2D region-based matching that assumes the optical flow is constant over a defined region, for example [1], favouring normalised cross correlation (NCC) compared to other matching criteria, and optical flow to estimate tissue motion. Cohen and Dinstein [2] and Boukerroui et al. [3] use an alternative speckle matching measure (CD_2), that assumes the speckle patterns in ultrasound images can be represented by a multiplicative Rayleigh distributed noise.

In our recent work [4] accurate interframe displacements and motion trajectories of individually tracked blocks were reported, using hierarchical blocks and a multiple scale NCC similarity measure. Focusing on musculoskeletal ultrasound, in deeper body regions a general reduction in correlation as a result of increased speckle noise was observed, affecting the correlation measure. Here, by combining two matching measures, we aim to maintain accuracy in strong signal regions using the first measure NCC, with low correlation and a low speckle signal to noise ratio (SNR) indicating necessary re-tracking using the second measure CD_2.

In this work, we favour displacement estimation with displacement post-processing, rather than speckle filter preprocessing and then displacement estimation. Although much research has been aimed at removing speckle to enhance ultrasound image understanding, many schemes produce increasingly homogeneous regions. This is due to features that are the same scale as the speckle being eliminated [5] impeding local motion estimation. Filter performance tends to be measured by quantifying edges and boundaries, with speckle preservation and fluctuation reduction measured using the co-occurrence matrix and localised mean and standard deviation (speckle SNR). In our situation all echo information is maintained, justifying a region-based motion estimation approach that has some inherent robustness to speckle incoherence and machine noise for speckle tracking.

Although substantial research exists using low frequencies at $3-7$ MHz (abdominal [2], cardiac and breast [3]), we focus on higher frequencies $8-16$ MHz for musculoskeletal diagnosis, capturing higher resolution images at a reduced penetration depth. This is due to attenuation where the signal is reduced by approximately 1 dB/cm/MHz [6]. We used three different probes (with bandwidths $5-10$, $8-16$ and $10-22$ MHz), to capture "perfect" conditions of an *in vitro* tendon section in a still water bath with clamped probe, and normal conditions of an *in vivo* freehand scanning of muscle. Sequences captured with perfect conditions were temporally stable resulting in high

*Corresponding Author: revell@cs.bris.ac.uk {*http://www.cs.bris.ac.uk/home/revell*}.

tracking accuracy using a single NCC tracking scheme. Sequences captured with normal conditions highlighted a reduction in correlation in areas of speckle, tending to occur in the lower regions, Fig. 1.

In the next section we concisely describe our datasets of simulated ultrasound speckle and *in vitro* tendons. Section 3 explains the combined matching measure approach, also defining displacement post-processing and interframe error measurement. Section 4 demonstrates and discusses sample interframe and trajectory results.

2 Ultrasound Datasets

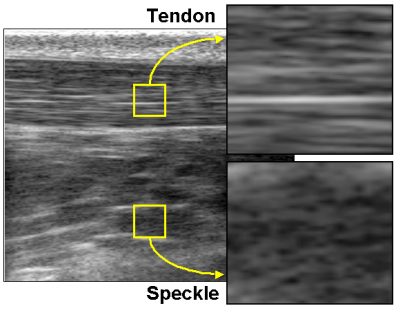
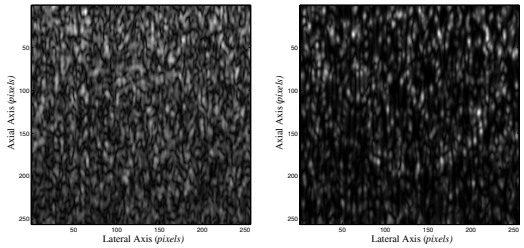


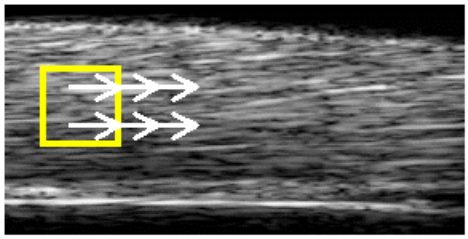
Figure 1. Achilles tendon with enlarged images of tendon and speckle.

To evaluate the advantages of the proposed method we generated spatially uniform and temporally stable speckle textures simulating an echographic speckle sequence [7], illustrated in Figs. 2(a)-2(b). The point spread function (PSF) $H(x,y)$ is assumed to be a Gabor function and the scattering function $T(x,y)$ a normally distributed random field that represents the population of scatterers being imaged. Convolving with the PSF yields the resulting RF echo data $I(x,y) = H(x,y) \otimes T(x,y)$, with envelope detection producing the desired image of echo magnitude. Speckle density is varied generating sequences of high (100%) and low (20%) speckle and varying speckle (containing half of each of these). To measure robustness against speckle reduced temporal coherence, we corrupt $I(x,y)$ with multiplicative Rayleigh distributed noise $\hat{I}(x,y) = \eta_m I(x,y)$ where $\eta_m \sim R(I) = \frac{I}{\sigma^2} \exp^{-\left(I^2/2\sigma^2\right)}$ with a non-zero mean specified by the single distribution parameter σ.

Several *in vitro* sequences were captured using an equine tendon that was pulled 3, 6 and 10 mm at known rates and loads whilst continuously scanning using an $8-16$ MHz clamped probe, Fig. 2(c). All sequences consist of 30 frames (the default acquisition length) captured at ≈ 10 Hz and quantised into 8 bits. All cycles included a positive and negative pull, similar to *in vivo* extension to flexion motions.



(a) Speckle density (100%). (b) Speckle density (20%).



(c) Longitudinal tendon section.

Figure 2. Sample images of (a) high (100%), (b) low (20%) density speckle and (c) an *in vitro* frame.

Typically, all ultrasound sequences will contain varying amounts of speckle and regions of underlying signal in different quantities. Therefore, we first analyse *in vitro* tendon sequences to demonstrate the good tracking results from using the NCC measure. Second, we analyse simulated speckle sequences to demonstrate the improvement of our proposed method where the speckle density varies across an image. Further information, including *in vivo* experiments using data as shown in Fig. 1, is available online*.

3 Proposed Method

The first measure NCC, applied in [4], assumes an increased SNR from high frequencies and sparse scatterers as shown in the tendon region in Fig. 1. Although we have found the correlation typically high, as described above, speckle noise reduces matching, highlighting the necessity of a suitable second measure. It must also be stated that other causes of correlation reduction are a lack of signal (probe de-coupling or curvilinear tendons), or signal saturation (incorrect gain controls or bone), or minimal features, causing problems for any similarity measure.

To combat reduced NCC accuracy in regions of increased speckle noise, we propose the use of a secondary measure instead of the NCC, namely the CD_2 measure, introduced by Cohen and Dinstein [2]. Recently, Boukerroui et al. [3] showed that in regions of fully developed speckle CD_2 is a more precise measure than for example, NCC or mean square error (MSE). CD_2 assumes (to be matched) blocks x and y from frames f_t and f_{t+1} are corrupted by independent multiplicative Raleigh distributed noise, representing uniform dense speckle. Log-compression

transforms the multiplicative noise to additive, denoted $\tilde{I}' = \ln(I')$ and $\tilde{I} = \ln(I)$. Following [3], we maximise the CD_2 objective function, where i and j are block and pixel indexes in $M \times N$ blocks, defined as:

$$\text{CD}_2 = \sum_{j=1}^{MN} \{(\tilde{I'}_{i,j} - \tilde{I}_{i,j}) - \ln(\exp(2(\tilde{I'}_{i,j} - \tilde{I}_{i,j})) + 1)\} \tag{1}$$

We use the two measures with multiple block scales, applying the NCC as the primary matching measure due to its high accuracy in low speckle density regions. However, we require an appropriate means of determining the amount of local speckle present. For this we use the SNR given by the ratio of the mean I_μ to the standard deviation I_σ of those pixels contained within a local region I, defined as $\lambda = \frac{I_\mu}{I_\sigma}$. In an area of uniform dense speckle, Wagner et al. [8] have determined an expected SNR value of $\lambda = 1.91$. We verified this value with *in vivo* data using multiple scaled regions of a uniform area, located at the focal zone of the ultrasound beam, and computed the mean SNR. Results showed that the mean SNR converged at $\lambda \approx 2$, hence we use a tolerance, empirically derived at 25%, to ensure a reasonable speckle sensitivity for *in vivo* images where speckle is seldom uniform. Only regions of $M \times N$, where M,$N \geq 16$ were able to determine reliably that a region contained uniform speckle. Therefore, we propose to use and apply either the NCC or the CD_2 measures, determined by the SNR, where $\lambda = 1.91$, which implies the speckle density present in a region:

$$\text{measure} = \begin{cases} \text{NCC} & \text{if SNR} > 1.25\lambda \\ \text{CD}_2 & \text{otherwise} \end{cases} \tag{2}$$

The SNR increases with a low amount of speckle (reaching infinity for specular reflection), justifying the NCC measure. However, SNR decreases for high amounts of speckle, indicating that the same block should be re-tracked using the secondary CD_2 measure. This is evaluated using the associated reference and candidate blocks in a full search with the same extents as the primary NCC measure for the larger scales. The speckle SNR is used as an indicator of speckle content (rather than correlation), as typically featureless regions of uniform speckle produce high correlation coefficients with surrounding speckle. Unfortunately the SNR is also sensitive to other image components, for example, feature boundaries resulting in a low local SNR, therefore we also check to ensure the NCC peak correlation coefficient c_{max} is low. This approach of alternating specific speckle and signal similarity measures using SNR allows the proposed method to adapt to image content.

Once the combined matching method is applied we perform displacement post-processing. Spurious velocity vectors are inevitable from any tracking process and are not always obvious. Potential causes are from noise or artifacts where multiple block scales have insufficient encapsulated features. Using a coherence based post-processing algorithm, adaptive weighted vector median filter (WVMF) [9], vector displacements are smoothed if inconsistent with their dominant neighbours whilst preserving motion boundaries. Given n displacements inside a sliding window, the WVMF output is a median displacement vector $\boldsymbol{d}^m$ that minimises the cumulative weighted p-norm distance between an individual $\boldsymbol{d}_i$ and neighbouring $\boldsymbol{d}_j$ displacements. A displacement is substituted with $\boldsymbol{d}^m$ if the cumulative weighted p-norm distance between $\boldsymbol{d}^m$ and $\boldsymbol{d}_i$ is significant, expressed as:

$$\sum_{i=1}^{n} w_i \|\boldsymbol{d}^m - \boldsymbol{d}_i\|_p \;\leq\; \sum_{i=1}^{n} w_i \|\boldsymbol{d}_j - \boldsymbol{d}_i\|_p \qquad j = 1, 2, \ldots, n \tag{3}$$

For combined measures our weighting uses the mean of both the NCC and CD_2 measures, defined $w = 0.5[c_{max} + (1 - \|\text{CD}_2\|)]$, ranging between 0 and 1. This technique can be iterative and lends itself to both interframe and trajectory smoothing with low computation.

To quantify displacement accuracy the error between the correct groundtruth displacement $\boldsymbol{d}_{t,c} = (u_c, v_c)$ and the estimated displacement $\boldsymbol{d}_{t,e} = (u_e, v_e)$ is measured by the angular error [10] combining errors in magnitude and direction into a single value:

$$\psi = \cos^{-1}(\breve{\boldsymbol{d}}_c \cdot \breve{\boldsymbol{d}}_e) \tag{4}$$

where ψ is the angle between the correct spatiotemporal vector $\breve{\boldsymbol{d}}_c = \frac{(u_c, v_c, 1)^T}{\sqrt{u_c^2 + v_c^2 + 1}}$ and, the estimated spatiotemporal vector $\breve{\boldsymbol{d}}_e = \frac{(u_e, v_e, 1)^T}{\sqrt{u_e^2 + v_e^2 + 1}}$. Further, displacement fields are used for displaced frame differencing, quantifying the root mean square (RMS) error between a backward reconstructed frame and the actual next frame.

4 Displacement Results and Discussion

Trajectories, which quantify continuous temporal displacements in sequences, were estimated for an *in vitro* region of tendon. The absolute errors (AE) between the groundtruth and mean estimated trajectories are summarised in

Table 1. The maximum AE was noticeable near the end of each pull cycle, due to the clamped tendon not returning to its original resting state. The mean AE remained low using NCC and combined measures. The NCC consistently produced high correlations $> 91\%$, producing significantly more accurate displacements compared to the single CD_2 and MSE measures. As mentioned later, due to the lack of speckle, using combined measures proved to be only as good as the single NCC measure. Typically, for *in vivo* data the speckle SNR instigates the usage of the NCC in minimal speckled (tendon) regions and the alternative CD_2 measure in dense speckle areas.

Pull	MSE			NCC			CD_2			NCC/CD_2 Combined		
mm	Max	Mean	STD Dev	Max	Mean	STD Dev	Max	Mean	STD Dev	Max	Mean	STD Dev
[a]3	4.79	3.60	3.53	2.16	2.69	1.21	5.57	3.88	2.23	2.16	2.71	1.22
[a]6	8.97	3.78	4.46	2.62	2.45	1.39	5.32	4.58	2.13	2.63	2.44	1.37
[a]10	25.74	10.35	7.96	5.54	7.79	3.12	19.30	9.00	4.11	5.55	7.79	3.12

Table 1. Summary of trajectory absolute error for *in vitro* tendon data for 3 pulls.

[a]Settings: WVMF iterations $= 2$ and multiple block scales where $M \times N = \{64, 32, 16, 8\}$.

Sample results in Table 2 quantify a comparative analysis between our proposed NCC/CD_2 combined approach and CD_2, NCC, MSE measures. The key improvement from our combined approach is observed in sequences with regions of varying speckle density (from multiple objects for example tendon and tissue), whereby using the appropriate measure allows for local signal variation; this is important for *in vivo* analysis. For regions of homogeneous speckle (high or low), the best accuracy is only as good as the appropriate single measure.

Speckle Pattern	High Density Speckle (100%)			Low Density Speckle (20%)			Varying Density Speckle		
Measures	Mean°	STD Dev°	RMS	Mean°	STD Dev°	RMS	Mean°	STD Dev°	RMS
[b]NCC/CD_2 Combined	**7.13**	**6.46**	**9.22**	**13.78**	**15.85**	**6.56**	**9.62**	**10.62**	**7.48**
[b]CD_2	**7.13**	**6.47**	**9.22**	15.95	16.63	7.55	11.46	13.91	9.33
[b]NCC	7.39	6.60	9.23	**13.72**	**15.83**	**6.56**	11.66	13.42	8.10
[b]MSE	7.58	7.45	13.27	23.74	24.55	14.20	13.48	18.05	10.15

Table 2. Interframe velocity angular error and displaced frame difference RMS error.

[b]Settings: Affine deformation (4 pel) sequences without noise using WVMF iterations $= 0$ and a block scale where $M \times N = \{16\}$.

Consequently, in Table 3 we show further results specifically for varying density speckle. These results illustrate a marked improvement compared to using a single measure. For frame pairs with applied noise, the NCC measure produced increasing errors, resulting from a reduction in correlation ranging between $96.84\% - 99.99\%$ to $72.75\% - 78.15\%$ between best (no applied noise) and worst $\sigma = 0.8$ cases. WVMF noticeably improved all results from just 2 iterations using an 8 neighbourhood region, maintaining a relatively low velocity angular error, for example 7.06 with and 9.62 without. Similar results were obtained from testing over 100 frame pairs.

Speckle Pattern	*Uncorrupted*		$\eta_m \sim \sigma = 0.4$		$\eta_m \sim \sigma = 0.8$	
Measures	Mean°	STD Dev°	Mean°	STD Dev°	Mean°	STD Dev°
[c]NCC/CD_2 Combined	**7.06**	**6.09**	**7.41**	**7.02**	**11.10**	**12.44**
[c]CD_2	7.54	6.92	9.94	10.34	12.84	15.09
[c]NCC	7.64	6.09	8.53	9.12	12.10	13.25
[c]MSE	7.30	7.49	20.18	18.81	21.28	20.09

Table 3. Interframe velocity angular error for 3 cases of noise corrupted frames of varying density speckle.

[c]Settings: Affine deformation (4 pel) sequences with varying noise using WVMF iterations $= 2$ and a block scale where $M \times N = \{16\}$.

We have demonstrated that using a combination of speckle pattern similarity measures improved interframe and trajectory performance, validating our approach on synthetic speckle and *in vitro* datasets. The real improvement in displacement accuracy is obvious from analysing frames that contain subregions ranging from dense speckle with noise characteristics that are purely multiplicative Rayleigh, to sparse stable speckle, to minimal speckle mixed with a strong underlying signal, all features typically found in *in vivo* data.

References

1. C. Wang, J.-M. Deng, G. Ateshian et al. "An automated approach for direct measurement of two-dimensional strain distributions within articular cartilage under unconfined compression." *Biomech. Eng.* **124**, pp. 557–567, 2002.
2. B. Cohen & I. Dinstein. "New maximum likelihood motion estimation schemes for noisy ultrasound images." In *16th International Conference on Pattern Recognition*, pp. 455–463. IEEE CS-Press, 2002.
3. D. Boukerroui, A. Noble & M. Brady. "Velocity estimation in ultrasound images: A block matching approach." In *18th IPMI*, volume 2732 of *LNCS*, pp. 586–598. Springer-Verlag, 2003.
4. J. Revell, M. Mirmehdi & D. McNally. "Strain quantification in ultrasound sequences." In *14th BMVC*, pp. 359–368. BMVA Press, 2003.
5. Y. Yu & S. Acton. "Speckle reducing anisotropic diffusion." *TIP* **11(11)**, pp. 1260–1270, 2002.
6. J. Thijssen & B. Oosterveld. "Texture in tissue echograms: Speckle or information?" *American Institute of Ultrasound in Medicine* **9**, pp. 215–229, 1990.
7. J. Meunier & M. Bertrand. "Ultrasonic texture motion analysis theory and simulation." *TMI* **14(2)**, pp. 293–300, 1995.
8. R. Wagner, S. Smith & J. Sandrik. "Statistics of speckle in ultrasound B-scans." *Soni. & Ultras.* **30(3)**, pp. 156–163, 1983.
9. L. Alparone, M. Barni, F. Bartolini et al. "Adaptively weighted vector-median filters for motion-field smoothing." In *IEEE International Conference on Acoustic Speech and Signal Processing*. 1996.
10. J. Barron, D. Fleet & S. Beauchemin. "Performance of optical flow techniques." In *12th JCCV*, pp. 43–77. 1994.

Comparison of Segmentation Methods for Cytometric Assay

Christophe Restif* and William F. Clocksin

Department of Computing, Oxford Brookes University, Wheatley, Oxford OX33 1HX, UK

Abstract. In this paper we compare two segmentation techniques, a parameter-free variant of the watershed algorithm and a new robust method based on adaptive contours. To test them, we use a set of fluorescence in-situ hybridisation images of leucocyte nuclei: these images were obtained in realistic conditions of mass screening, and do not meet the prerequisites of usual segmentation methods. We present the two methods and describe them step by step. We then compare the results of the segmentation of the test images, showing over a hundred nuclei, and detail the incorrect segmentations. The contour-based method outperforms the watershed as regards correct segmentations and over-segmentations, but performs more under-segmentations. We analyse the origins of the incorrect segmentations from the two methods, and discuss possible improvements.

1 Introduction

We are interested in segmenting images of human leucocyte nuclei, obtained by fluorescence-in-situ hybridisation (FISH). On such images, nuclei are visible in the blue channel as bright circles on a darker background, while specifically marked parts of the chromosomes – such as telomeres – appear as bright dots in the green or red channels. A typical image is shown on Figure 1. Counting the number of dots inside each nucleus or measuring their intensities is a key step in identifying certain genetic diseases, or in differentiating nuclei from different individuals: thus, it is crucial to locate the nuclei boundaries accurately. This prerequisite operation is called segmentation: it can be done manually with a drawing tool, semi-automatically by clicking on nuclei to be bounded, or automatically, possibly requiring the tuning of parameters. In this paper we focus on the segmentation task, and consider only the blue channel of images, as grey-scale images.

Of all the segmentation techniques that have been developed, many are tailored for a specific type of cell under specific imaging conditions, as in [1, 2]. However, our images do not meet the requirements of most of these techniques, as they were obtained from a prototype of a mass screening system: in particular the backgrounds are far from uniform, and nuclei tend to gather in small clusters. Manual segmentation is a trivial task, but is too time-consuming for mass screening. Thus, we investigated two popular methods with a few variants that have been used for a wide variety of cells: one is based on the watershed transform, the other uses contours. General reviews of these methods can be found in [3] and [4] respectively. In this paper we describe the methods briefly, then present the results we obtained on a set of FISH images; finally we discuss their inherent limitations and some possible enhancements.

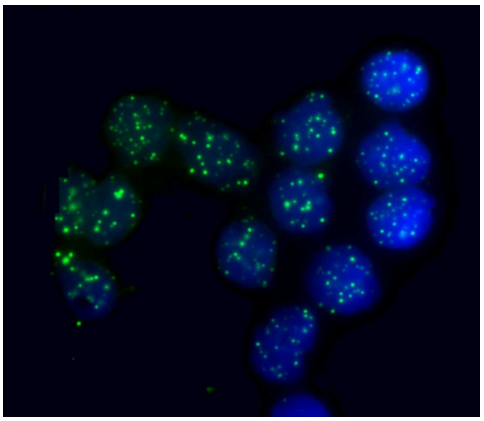


Figure 1. FISH image showing leucocyte nuclei: the bright dots within the nuclei are the telomeres.

2 Segmentation methods

2.1 Watershed-based segmentation

The watershed transform is a method used to segment a grey-scale image into zones of influence of local minima. Intuitively, one can think of a grey-scale image as a landscape, with dark regions at low altitude and bright regions at high altitude. Supposing that every single depression in this landscape is pierced at the bottom, and that water starts rising from below, then gradually basins will appear and fill up, until eventually every valley and mountain is immersed. The idea is to mark the lines where the basins actually merged during the flooding: these lines will divide the landscape into regions called 'catchment basins'. The resulting segmented landscape is called the watershed transform of the image: it shows the boundaries of the regions of interest. Proper definitions of the watershed transform, the watershed algorithms and their implementations are presented in details in [3].

*Corresponding author. E-mail: crestif@brookes.ac.uk

By definition, the transform will create one catchment basin for every local grey minimum: that generally results in an extreme over-segmentation of the image. Thus, the key point of any watershed algorithm is to use some preprocessing to ensure that one and only one local minimum is to be found within each region of interest – and that none will be found outside. In our case, every catchment basin should correspond to a nucleus: in other words, the preprocessing needs to locate one local minimum within each nucleus.
Preprocessing usually involves basic image morphological operations, such as thresholding, opening, closing, etc. We implemented the watershed segmentation as described in [5]. It presents a variety of preprocessing steps, illustrated on two different types of nuclei, namely bone marrow (BM) and peripheral blood (PB) nuclei. The very first steps of the segmentation, namely the image acquisition and the shading correction, are specific to the type of imaging hardware used, which is why we do not detail them here. The remaining parts of the preprocessing consist in a thresholding followed by an image transformation, specific to each type of nucleus.
Thresholding turns a grey image into black and white. A wide range of thresholding techniques is available [6]. The one used in [5] is the isodata algorithm: the threshold value is found iteratively, in such a way that it divides the histogram into two regions, whose two mean values are equidistant to the threshold. Its main strength is that it is automatically computed for each image. Besides this thresholding, a method is mentioned to isolate internuclei background: h-domes are extracted from the original image, so that with a carefully chosen parameter h, the internuclei background appears as the biggest bright parts of the extraction. These can be isolated with a standard top-hat transform, and added as background to the thresholded image. The h-dome extraction is detailed in [7].
The final part of the preprocessing depends on the type of nucleus under observation. As PB nuclei appear almost round, even when packed in clusters, the thresholded image from the previous step will show a juxtaposition of white circles on a dark background. The distance transform is applied on this image: each white pixel is turned into a grey pixel, whose brightness is proportional to the distance of the nearest black pixel. The result is a grey picture, where the central regions of the circles appear brighter, and their borders gradually merge with the black background. As for BM nuclei, such a distribution of grey levels within the nuclei is observed in the original image: the brightness decreases smoothly from the centre of a nucleus to its border. Enhancing the thresholded image with the original grey values is thus enough for this step.
This preprocessing removes most of the unwanted local minima (ideally all of them); the remaining minima can then be marked as starting points – 'pierced' – for the flooding process. One way to locate them is to use h-dome extractions, with a different value of h for each type of nucleus: h-domes merge minima that are very close to each other, thus reducing over-segmentation. Then the watershed transform is performed. The core watershed algorithm in use is the one described in [8]; a more detailed version can be found in [3].

2.2 Contour-based segmentation

The second segmentation method we implemented is based on contours. A review of active contour-based segmentation of medical images can be found in [4]. All the details of the algorithm we used are presented in [9]. In short, it is a two-step algorithm, consisting in localising the nuclei and then their boundaries. During the first step, a single point is found within each nucleus, using a variant of the mean-shift algorithm [10]. In the second step, contours are grown iteratively from these seed points, until they fit the nuclei boundaries. Contours are deformable and remeshable polygons, whose expansions and contractions are controlled by an energy function, computed at every vertex and every iteration. Internal energy controls the general shape of the polygon regardless of the image, while external energy drives the vertices individually towards the boundaries of the nuclei. In addition, a repulsive energy ensures that two contours will not grow into each other: this is needed when segmenting images containing clusters of nuclei. Overall, the seed finding and contour fitting make use of a dozen parameters. They have to be tuned manually on test images.

3 Results

To test the two methods, we used a set of 23 FISH images of human leucocyte nuclei, showing a total of 111 nuclei. Each image contained from one to sixteen nuclei, either isolated (48 nuclei) or packed together in clusters (63 nuclei).

We implemented the watershed algorithm with the two preprocessing steps described above. However, using the original values for the watershed transform failed on every image that we tested, resulting in an over-segmentation of the nuclei. This was to be expected, since none of them displayed uniform radial intensity variation, unlike PB nuclei. Table 1 shows the results of the segmentation after the distance transform preprocessing.

As explained above, the watershed transform will create as many segmented regions as there are local minima in

number of nuclei	analysed	correctly segmented	over-segmented	fused	ignored
isolated nuclei	48	44 (92%)	4 (8%)	0 (0%)	0 (0%)
nuclei in clusters	63	42 (67%)	6 (10%)	4 (6%)	11 (17%)
all cases	111	86 (77%)	10 (9%)	4 (4%)	11 (10%)

Table 1. Watershed segmentation results

the first place. Thus, the final segmentation results depend only on the relevance of the threshold values found by the isodata algorithm, and on the minima created by the distance transform and located by h-domes. The nuclei ignored had a grey intensity lower than the threshold. Over-segmentation happens when two or more minima were located within one nucleus, and fusion happens when no local minimum was located within a nucleus in a cluster.

To test the contour-based segmentation, we set up the parameters manually on a subset of images: these parameters were then kept for the analysis of the whole set of test images. The results of the contour-based segmentation are listed in Table 2.

number of nuclei	analysed	correctly segmented	over-segmented	fused	ignored
isolated nuclei	48	45 (94%)	2 (4%)	0 (0%)	1 (2%)
nuclei in clusters	63	48 (76%)	3 (5%)	12 (19%)	0 (0%)
all cases	111	93 (84%)	5 (4%)	12 (11%)	1 (1%)

Table 2. Contour segmentation results

It is to be noted that all the incorrect segmentations result from the first step, the seed finding. Over-segmentation occurs when two or more seeds are found within one nucleus; fusion of clustered nuclei happens when the algorithm finds fewer seeds in a cluster than there are nuclei. Finally, a nucleus is ignored when no seed is found inside it. Once the seeds are found, the contour-fitting step adjusts the boundaries of the nuclei, but does not modify the segmentation results in one way or another.

4 Discussion

We identified two intrinsic problems with this version of the watershed when applied to our set of images. First, even though the distance transform can create a single minimum within an isolated nucleus or a nucleus in a not-closely packed cluster, it fails when nuclei are closely packed, resulting in an under-segmentation of the cluster. Also, it is based on the assumption that nuclei are round, and thus tend to over-segment irregular shapes – mistaken for clusters. Over-segmentation can be corrected with tailored post-processing, such as hierarchical or multiscale watershed [11, 12]: using the landscape metaphor, the idea is to measure the minimum height that a point in a local minimum has to climb before reaching a lower local minimum. If that height is smaller than a certain threshold, then the original minima can be thought of as shallow, and its attraction basin can be considered as an artifact – as opposed to a nucleus – and merged with its lower neighbour. This method is computationally intensive, and requires the tuning of one parameter: the threshold value differentiating shallow minima from deep ones. There is little that can be done concerning the under-segmentation of clusters though.
The second problem comes from the thresholding itself. Some images showing a cluster of a dozen nuclei have a non-uniform background that cannot be smoothed easily. In one image, the general intensity in the top part is much higher than in the bottom part, and as a result some parts of the background are brighter than some parts of nuclei. In another image, the background near the nuclei is very dark, while the background further away is brighter, even brighter than some nuclei. In these cases, any threshold will either cut off some of the nuclei, or merge others with the background. Alternative methods to the isodata algorithm provide a threshold surface, such as [13], thus allowing variations of the threshold value across the image. However, they also require the tuning of parameters, thus loosing the automation provided by the isodata algorithm.
Overall, the main disadvantage of the watershed segmentation is that it removes most of the information contained in the original image. In particular, the slight intra-cluster variations of intensities are crucial for a human eye to segment a cluster properly, but are erased by the thresholding. The following steps, such as the dome extraction or the distance transform, tend to recreate an artificial basic version of the original image, but are bound to be specific to every type of nucleus, and not easy to automate.

Conversely, the contour-based method uses the original image at all time. Judging from our experiments in parameter tuning, it appears that only one parameter is critical: the range of attraction of candidate points in the mean-shift-based seed-finding step [9], which corresponds directly to the typical size of a nucleus on the current image. In the tests reported here, that parameter was fixed, although the sizes of nuclei varied from one image to another, by up to two-fold. This explains the high number of fusions reported in Table 2. Some methods have been developed to automate the tuning of that parameter [14]. Combined with quality control methods, as in [15], they appear as a promising improvement towards automatic segmentation. Basically, the stability of the segmentation is tested against the parameter to control: the whole range of values is used, and all the corresponding segmentation results are compared. When these results are consistent over a range of values, they are considered as reliable; otherwise the image is marked as requiring further care – such as manual segmentation. This has to be done for each parameter separately. Thus, it seems to be restricted to segmentation methods with few parameters. Besides, it requires a preliminary study to prove the link between the stability of segmentation results and their reliability.

5 Conclusion

In our experiment, the contour-based method outperforms the watershed algorithm, yet both methods could be improved with some post-processing of the results. However, the consequences of these improvements are different for each method. Introducing parameters to perform non-uniform thresholding or multi-scale watershed removes the main strength of the watershed algorithm; whereas testing the stability of the contour-based segmentation is a further step to improve the quality control of the overall method.

Acknowledgements

The authors wish to thank Maj Hulten, from the University of Warwick, UK, for providing the FISH images.

References

1. T. Tanaka, T. Joke & T. Oka. "Cell nucleus segmentation of skin tumor using image processing." In IEEE (editor), *2001 Proceedings of the 23rd Anual EMBS International Conference*, volume 3, pp. 2716–2719. October 2001.
2. C. Ortiz de Solorzano, R. Malladi, S. Lelievre et al. "Segmentation of nuclei and cells using membrane related protein markers." *Journal of Microscopy* **201(3)**, pp. 404–415, March 2001.
3. J. B. Roerdink & A. Meijster. "The watershed transform: Definitions, algorithms and parallelization strategies." *Fundamenta Informaticae* **41**, pp. 187–228, 2001.
4. T. Lehmann, J. Bredno & K. Spitzer. "On the design of active contours for medical image segmentation." *Methods of Information in Medicine* **42(1)**, pp. 89–98, 2003.
5. N. Malpica, C. Ortiz de Solorzano, J. J. Vaquero et al. "Applying watershed algorithms to the segmentation of clustered nuclei." *Cytometry* **28**, pp. 289–297, 1997.
6. M. Sezgin & B. Sankur. "Survey over image thresholding techniques and quantitative performance evaluation." *Journal of Electronic Imaging* **13(1)**, pp. 146–168, January 2004.
7. L. Vincent. "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms." *IEEE Transactions on Image Processing* **2(2)**, pp. 176–201, April 1993.
8. L. Vincent & P. Soille. "Watersheds in digital spaces: An efficient algorithm based on immersion simulations." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13(6)**, pp. 583–598, June 1991.
9. W. Clocksin. "Automatic segmentation of overlapping nuclei with high background variation using robust estimation and flexible contour models." In *Proceedings of the 12th International Conference on Image Analysis and Processing*, pp. 682–687. Mantova, Italy, September 2003.
10. D. Comaniciu & P. Meer. "Mean shift: A robust approach toward feature space analysis." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24(5)**, pp. 603–619, May 2002.
11. S. Beucher. "The watershed transformation applied to image segmentation." *Scanning Microscopy* **6**, pp. 299–314, 1992.
12. E. Dougherty & R. Lotufo. *Hands-on Morphological Image Processing - Chapter 7: Morphological Segmentation - Watershed*, volume TT59, pp. 139–154. SPIE – The International Society for Optical Engineering, July 2003.
13. F. H. Chan, F. Lam & H. Zhu. "Adaptive thresholding by variational method." *IEEE Transactions on Image Processing* **7(3)**, pp. 468–473, March 1998.
14. D. Comaniciu. "An algorithm for data-driven bandwidth selection." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25(2)**, pp. 281–288, February 2003.
15. P. Bamford & B. Lovell. "Method for accurate unsupervised cell nucleus segmentation." In *Proceedings IEEE Engineering in Medicine and Biology*. 2001.

Spatio-Temporal Free-Form Registration of Cardiac MR Image Sequences

D Perperidis[a,*], R. Mohiaddin[b], D. Rueckert[a]

[a]Visual Information Processing Group, Department of Computing, Imperial College, London,
[b] Cardiovascular MR Unit, Royal Brompton Hospital, Imperial College, London

Abstract. In this paper we develop a spatio-temporal registration algorithm for cardiac MR image sequences. The algorithm has the ability to correct any spatial misalignment between the images caused by global differences in the acquisition of the image sequences and by local shape differences. In addition it has the ability to correct temporal misalignment caused by differences in the length of the cardiac cycles and by differences in the dynamic properties of the hearts. The algorithm uses a 4D deformable transformation model which is separated into spatial and temporal components. The registration method was qualitatively evaluated by visual inspection and by measuring the overlap and surface distance of anatomical regions. The results demonstrate that a significant improvement in the alignment of the image sequences is achieved by the use of the deformable transformation model.

1 Introduction

Cardiovascular diseases are a very important cause of death in the developed world. Their early diagnosis and treatment is crucial in order to reduce mortality and to improve patients' quality of life. Magnetic resonance imaging (MR) is playing an increasingly important role for the high resolution imaging of the cardiovascular system. The recent advantages in the development of cardiac imaging modalities have led to an increased need for cardiac registration methods. In general, cardiac image registration is a very complex problem due to the complicated non-rigid motion of the heart and the thorax as well as the low resolution with which cardiac images are usually acquired. In the recent years cardiac image registration has emerged as an important tool for a large number of applications. It has a fundamental role in the construction of anatomical atlases of the heart [1], analysis of the myocardial motion [2] and in fusion of information from a number of different modalities such as CT, MR, PET, and SPECT [3, 4]. In addition, cardiac image registration is crucial for the comparison of images of the same subject, e.g. before and after pharmacological treatment or surgical intervention. Furthermore, inter-subject alignment of cardiac image sequences to the same coordinate space (anatomical reference) enables direct comparisons between the cardiac anatomy and function of different subjects to be made.

While a large number of registration techniques exist for cardiac images, most of these techniques focus on 3D images ignoring any temporal misalignment between the two image sequences. In this paper we extend a 4D cardiac MR image registration method based on voxel similarity measures which has been recently presented [1,5,6]. This method will not only bring a number of sequences of cardiac images acquired from different subjects or the same subject (for example short and long axis cardiac image sequences) into the same spatial coordinate frame but also into the same temporal coordinate frame. The aim of this contribution is to improve the accuracy of the cardiac MR image sequence registration algorithm by using a spatio-temporal free-form deformation model based on B-Splines. The algorithm has the ability to correct any spatial misalignment between the images caused by global differences in the acquisition of the image sequences and by local shape differences. In addition it has the ability to correct temporal misalignment caused by differences in the length of the cardiac cycles and by differences in the dynamic properties of the hearts.

2 Spatio-Temporal Registration

Since the heart is undergoing a spatially and temporally varying degree motion during the cardiac cycle, 4D cardiac image registration algorithms are required when registering two cardiac MR image sequences. Spatial alignment of corresponding frames of the image sequences (e.g. the second frame of one image sequence with the second frame of the other) is not enough since these frames may not correspond to the same position in the cardiac cycle of the hearts. This is due to differences in the acquisition parameters, differences in the length of cardiac cycles and differences in the dynamic properties of the hearts. Spatio-temporal alignment will enable comparison between corresponding anatomical positions and corresponding positions in the cardiac cycle of the hearts.

*email:dp1@doc.ic.ac.uk

A 4D cardiac image sequence can be represented as sequence of n 3D images $I_k(x, y, z)$ with a fixed field of view Ω_I and an acquisition time t_k with $t_k < t_{k+1}$, in the temporal direction. The resulting image sequence can be viewed as 4D image $I(x, y, z, t)$ defined on the spatio-temporal domain $\Omega_I \times [t_1, t_n]$. The goal of 4D image registration described in this paper is to relate each point of one image sequence to its corresponding point of the reference image sequence:

$$\mathbf{T}(x, y, z, t) = (x'(x, y, z), y'(x, y, z), z'(x, y, z), t'(t)) \tag{1}$$

and can be of a subvoxel displacement in the spatial domain and of a sub-frame displacement in the temporal domain. The 4D mapping can be resolved into decoupled spatial and temporal components $\mathbf{T}_{spatial}$ and $\mathbf{T}_{temporal}$ respectively where

$$\mathbf{T}_{spatial}(x, y, z) = (x'(x, y, z), y'(x, y, z), z'(x, y, z)), \mathbf{T}_{temporal}(t) = t'(t)$$

One consequence of this decoupling is that each temporal frame t in image sequence I will map to another temporal frame t' in image sequence I', ensuring causality and preventing different regions in a 3D image $I_t(x, y, z)$ from being warped differently in the temporal direction by $\mathbf{T}_{temporal}$.

2.1 Spatial Alignment

The aim of the spatial part of the transformation is to relate each spatial point of a particular image of one image sequence to a point in another particular frame of the reference image sequence. The transformation $\mathbf{T}_{spatial}$ consists of a global transformation and a local transformation:

$$\mathbf{T}_{spatial}(x, y, z) = \mathbf{T}^{global}_{spatial}(x, y, z) + \mathbf{T}^{local}_{spatial}(x, y, z) \tag{2}$$

The global transformation is an affine with 12 degrees of freedom transformation addressing differences in the size, orientation and alignment of the hearts while the local part addresses differences in the shape of the hearts. A free-form deformation (FFD) model based on B-splines is used in order to describe the differences in the local shape of the hearts. To define a spline-based FFD we denote the spatial domain of the image volume as $\Omega_I = \{(x, y, z) \mid 0 \leq x < X, 0 \leq y < Y, 0 \leq z < Z\}$. Let Φ denote a $n_x \times n_y \times n_z$ mesh of control points $\phi_{i,j,k}$ with uniform spacing δ. Then, the FFD can be written as the 3D tensor product of the familiar 1D cubic B-splines [7]:

$$\mathbf{T}^{local}_{spatial}(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(w) \phi_{i+l, j+m, k+n} \tag{3}$$

One advantage of B-Splines is that they are locally controlled which makes them computationally efficient even for a large number of control points.

2.2 Temporal Alignment

The temporal part of the transformation consists of a temporal global part, $\mathbf{T}^{global}_{temporal}$, and a temporal local part, $\mathbf{T}^{local}_{temporal}$:

$$\mathbf{T}_{temporal}(t) = \mathbf{T}^{global}_{temporal}(t') + \mathbf{T}^{local}_{temporal}(t')$$

$\mathbf{T}^{global}_{temporal}$ is an affine transformation which corrects for differences in the length of the cardiac cycles and differences in the acquisition parameters. $\mathbf{T}^{local}_{temporal}$ is modeled by a free-form deformation using a 1D B-spline and corrects for temporal misalignment caused by different cardiac dynamic properties (differences in the length of contraction and relaxation phases, different motion patterns, etc). To define a spline-based temporal free-form deformation we denote the temporal domain of the image sequence as $\Omega_t = \{(t) \mid 0 \leq x < T\}$. Let Φ denote a set of n_t control points ϕ_t with a temporal spacing δ_t. Then, the temporal free-form deformation can be defined as a 1D cubic B-spline:

$$\mathbf{T}^{local}_{temporal}(t) = \sum_{l=0}^{3} B_l(u) \phi_{t_{i+l}} \tag{4}$$

where $i = \lfloor \frac{t}{n_t} \rfloor - 1$, $u = \frac{t}{n_t} - \lfloor \frac{t}{n_t} \rfloor$ and B_l represents the l-th basis function of the B-spline.

$\mathbf{T}^{local}_{temporal}$ deforms the temporal characteristics of each image sequence in order to follow the same motion pattern with the reference image sequence. The combined 4D transformation model (equation 1) is the spatio-temporal

free-form deformation (STFFD). The optimal transformation is found by maximising a voxel based similarity measure, the Normalised Mutual Information (NMI), calculated directly from the the joint intensity histogram of the two sequences over the spatio-temporal domain of overlap. In the first part of the optimization procedure NMI is optimized as a function of$\mathbf{T}_{spatial}^{global}$ and $\mathbf{T}_{temporal}^{global}$ using an iterative downhill descent algorithm. In the second part, NMI is optimized as a a function of $\mathbf{T}_{spatial}^{local}$ and $\mathbf{T}_{temporal}^{local}$ using a simple iterative gradient descent method.

3 Results and Discussion

To evaluate the spatio-temporal deformable registration algorithm we have acquired fifteen cardiac MR image sequences from healthy volunteers using a Siemens Sonata 1.5 T scanner and TrueFisp pulse sequence. For the reference subject 32 different time frames were acquired (cardiac cycle of length 950msec). Each 3D image of the sequence had a resolution of $256 \times 192 \times 46$ with a pixel size of $0.97mm \times 0.97mm$ and a slice thickness of 3mm. Fourteen 4D cardiac MR images were registered to the reference subject. An initial estimate of the global spatial transformation was provided due to the large variety in the position and orientation of the hearts. Since all the image sequences contained almost entire cardiac cycles, the global temporal transformation was calculated in order to compensate the differences in length of the cardiac cycles of the subjects (by matching the first and the last frames of the image sequences). The spacing of the control points of the local transformation were 10mm in the spatial domain and 90msec in the temporal domain.

Figure 1 provides an example of the spatio-temporal free-form registration. The images in the top row (a-c) are the short axis, the long axis and the temporal views of a frame in the middle of the image sequence after the optimization of the global transformations (affine spatio-temporal registration). The lines in the images represent the contours of the reference image sequence. The images in the bottom row of figure 1 are the same images after spatio-temporal free-form registration. The quality of the registration in the spatial domain was measured by calculating the volume overlap for the left and right ventricles as well as for the myocardium. The volume overlap for an object O is defined as:

$$\Delta(T,S) = \frac{2 \times |T \bigcap S|}{|T| + |S|} \times 100\% \tag{5}$$

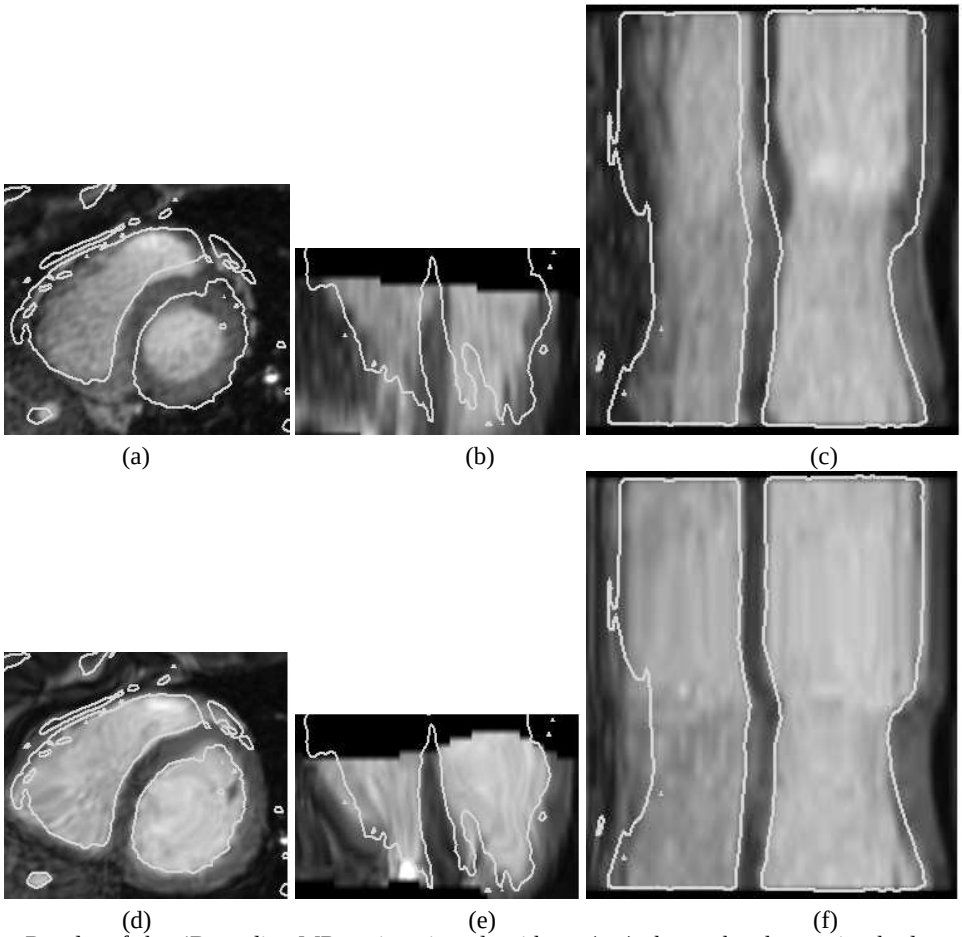
Here T denotes the voxels in the reference (target) image part of object O and S denotes the voxels in the other image part of object O. We have also calculated the mean surface distance of the above anatomical regions after the affine and the deformable 4D registration. In order to calculate the overlap of the anatomical structures and the surface distance we used segmented images. Table 1 shows the mean volume overlap and the mean surface distance (in mm) for each anatomical region after spatio-temporal affine registration, after 3D non-rigid registration of the first frames (by matching the first and the last time frames of the image sequences) and after the spatio-temporal deformable registration.

	Volume Overlap			**Surface Distance in mm**		
Anatomical region	**Affine 4D**	**Deformable 3D**	**Deformable 4D**	**Affine 4D**	**Deformable 3D**	**Deformable 4D**
Left ventricle	76.15%	80.95%	85.76%	4.16	3.56	2.96
Right ventricle	77.39%	83.87%	84.86%	4.95	4.27	3.60
Myocardium	70.39%	71.64%	73.86%	4.77	3.87	4.16

Table 1. The mean volume overlap and surface distance after the affine 4D registration,after the deformable 3D and after spatio-temporal deformable registration.

4 Conclusions and Further Work

We have presented an extension to our earlier spatio-temporal registration method [5, 6]. In this contribution we extended the registration method by introducing a spatio-temporal deformable transformation model. The proposed registration approach corrects temporal misalignment caused by different acquisition parameters, different length of cardiac cycles and different dynamic properties of the hearts. The approach also corrects spatial misalignment caused by global differences in the acquisition of the image sequences and local shape differences. There is a large range of applications for this spatio-temporal registration method. We are planning to use it for building a probabilistic atlas of the cardiac anatomy and function similar to the one we have recently presented [1]. The method was evaluated using fifteen image sequences from healthy volunteers. The results indicate a significant improvement in the temporal and spatial registration of the image sequences with the use of the spatio-temporal deformable transformation model.



(a) (b) (c)

(d) (e) (f)

Figure 1. Results of the 4D cardiac MR registration algorithm. (a-c) shows the short axis, the long axis and the temporal views after the affine alignment, (d-f) shows the corresponding short axis, long axis and temporal views after the spatio-temporal free-form registration. Animations of the registrations can be found at `http://www.doc.ic.ac.uk/~dp1/Conferences/MIUA04/REGISTRATION/`

References

1. D. Perperidis, M. Lorenzo-Valdés, R. Chandrashekara, A. Rao, R. Mohiaddin, G.I Sanchez-Ortiz, and D. Rueckert. Building a 4D atlas of the cardiac anatomy and motion using MR imaging. In *2004 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2004.
2. R. Chandrashekara, R. H. Mohiaddin, and D. Rueckert. Analysis of myocardial motion in tagged MR images using nonrigid image registration. In M. Sonka and J. Michael Fitzpatrick, editors, *Proceedings of the SPIE International Symposium on Medical Imaging*, pages 1168–1179, San Diego, California USA, 24–28 February 2002.
3. M.C. Gilardi, G. Rizzo, A. Savi, C. Landoni, V. Bettinardi, C. Rosseti, G. Striano, and F. Fazio. Correlation of SPECT and PET cardiac images by a surface matching registration technique. *Computerized Medical Imaging and Graphics*, 22(5):391–398, 1998.
4. T. G. Turkington, T. R. DeGradom, M. W. Hanson, and E. R. Coleman. Alignment of dynamic cardiac PET images for correction of motion. *IEEE Transaction on Nuclear Science*, 44(2):235–242, 1997.
5. D. Perperidis, A. Rao, M. Lorenzo-Valdés, R. Mohiaddin, and D. Rueckert. Spatio-temporal alignment of 4D cardiac MR images. In *Lecture Notes in Computer Science: Functional Imaging and Modeling of the heart, FIMH'03*, Lyon, France, June 5-6, 2003. Springer.
6. D. Perperidis, A. Rao, R. Mohiaddin, and R. Rueckert. Non-rigid spatio temporal alignment of 4D cardiac MR images. In *Lecture Notes in Computer Science: Second International Workshop on Biomedical Image Registration (WBIR)*. Springer, 2003.
7. D. Rueckert, L. I. Sonoda, C. Hayes, D.L.G Hill, M.O. Leach, and D.J Hawkes. Non-rigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.
8. C. Studholme, D.L.G. Hill, and D.J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1999.
9. M. Lorenzo-Valdés, G. I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert. Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. In *Sixth Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI'03)*, Lecture Notes in Computer Science. Springer-Verlag, 2003.

Stereo-photogrammetry and image analysis for Fetal Alcohol Syndrome screening

Tania S Douglas[a], Ernesta M Meintjes[a], Christopher L Vaughan[a], Rex Grobbelaar[a] and Denis L Viljoen[b]

[a]MRC/UCT Medical Imaging Research Unit, Dept of Human Biology, University of Cape Town, South Africa
[b]Dept of Human Genetics, University of the Witwatersrand, South Africa

1 Introduction

Fetal alcohol syndrome (FAS), a consequence of excessive maternal alcohol consumption during pregnancy, is one of the most common preventable causes of mental retardation worldwide. FAS prevalence of 40.5 to 46.4 per 1000 children aged 5 to 9 years has been reported in one disadvantaged community in South Africa, while the rate for the developed world has been estimated to range from 0.33 to 2.2 per 1000 [1]. FAS is therefore a serious public health problem in South Africa. Large epidemiological studies are required in this country to identify populations that are especially at risk, so that prevention and intervention programmes can be introduced. To this end, a cost-effective and easily deployable FAS screening tool is required.

FAS diagnosis is made on clinical grounds alone, with an emphasis on its unique facial phenotype, which is assessed by distance measurements between landmarks around the eyes and mouth [2,3]. Conventional diagnosis of the facial phenotype requires intrusive direct measurements on the face performed by trained specialists; these measurements are compared to normative data to determine whether subjects have FAS. Eye measurements used in FAS diagnosis are shown in figure 1. Direct measurements are intrusive to the patient, time-consuming and costly, and not conducive to large-scale screening for and surveillance of FAS. As an alternative, photogrammetry has been introduced as an indirect method of obtaining distances between landmarks on the face, using both two- and three-dimensional methods. The period of interaction with the patient is potentially shorter for indirect measurements, since features are measured after data acquisition. Indirect methods are therefore also less dependent on patient behaviour and the need for the patient to keep still for long periods, and have particular advantages when children are being examined. Some measurements, such as those around the eyes, are difficult to obtain directly without risk of discomfort or injury to the patient.

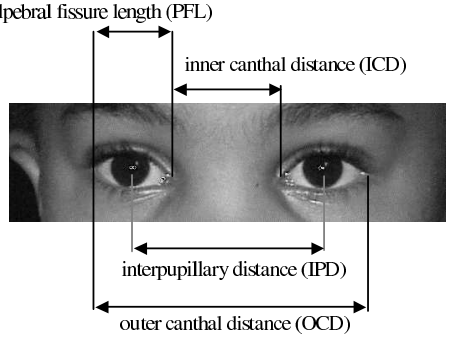


Figure 1. Distance measurements used in FAS diagnosis

We review our development of a stereo-photogrammetric instrument and image analysis algorithms, using established methods, to diagnose the facial phenotype associated with FAS in children.

2 Image acquisition

Our stereo-photogrammetric tool is shown in figure 2 and comprises 2 digital cameras, a chin- and headrest surrounded by calibration markers within which the child's face is photographed, and a notebook computer [4]. Stereo-photogrammetric algorithms, based on the Direct Linear Transform (DLT), are used to calculate the three-dimensional coordinates of and distances between facial landmarks selected by an operator from stereo

Email: tdouglas@cormack.uct.ac.za

digital images by mouse clicks on a computer monitor. The equipment fits into a robust wheeled carrying case for easy transportation.

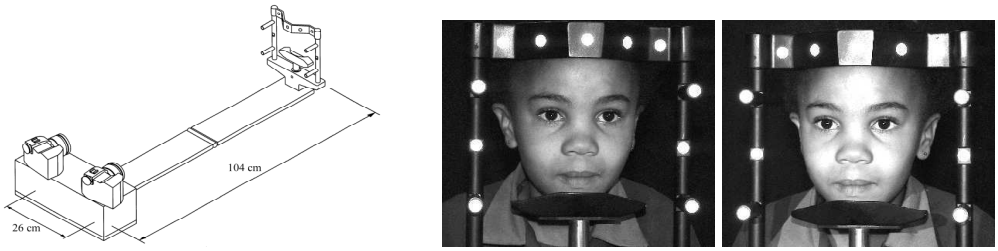


Figure 2. Image acquisition instrument for FAS screening and stereo images of a 6 year-old child

2.1 Comparison with single frontal photographs and direct measurements

We have compared measurements obtained using stereo-photogrammetric calibration with those obtained from single frontal photographs. We have found that, in an ideal system, where the real-world co-ordinates of the relevant points on the eyes are known from three-dimensional calibration, depth does not make a significant contribution to the distance measurements involving the eyes that are required for fetal alcohol syndrome screening. Two-dimensional measurements provide reliable results in such cases. However, in a real situation, using single frontal photographs, where the exact relationships of the camera with respect to the calibration frame is not known, and where the eyes are not placed at the same depth as the calibration markers, the accuracy of two-dimensional measurements is questionable [5].

We have also shown that eye measurements obtained using our stereo-photogrammetric instrument compare favourably with direct measurements obtained by clinicians [4].

3 Eye feature extraction

Image processing algorithms applied to facial photographs have the potential to enhance photogrammetry, through automatic extraction of desired facial features and the subsequent automatic measurement of clinically relevant distances, resulting in further reductions in the time spent on examinations and improvements in the reliability of measurements.

The points required to measure the eye distances shown in figure one, are the pupil centres and the inner and outer canthi (the eye corners). An algorithm was developed to extract these automatically [6].

Since the face is always photographed in a fixed position, an image window in which the eyes are expected to appear can be specified in order to reduce the image area being examined. Estimates of the eye centres are obtained by combining a peak and a valley map of the eye window and using integral projections of the combined map to identify areas of high intensity representing the eyes. Connected components extracted from the area surrounding the eye centre estimate in the peak map represent the sclera (figure 3).

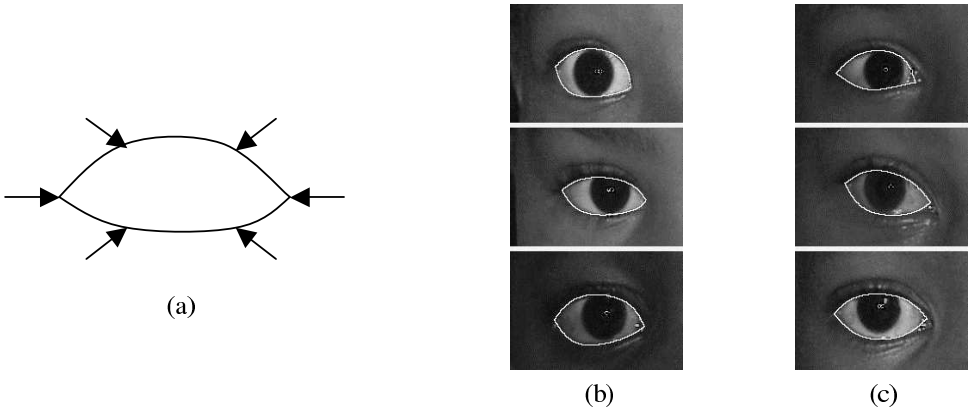


(a) (b)

Figure 3. Extracted (a) sclera region and (b) iris contour

The eye contour is found by using a genetic algorithm (GA) to match cubic splines to the valley map, in which the eye contour appears bright. Two four-point cubic splines with common endpoints, one for the upper and one for the lower eye contour, represent the eye template (figure 4 (a)). The parameters being optimised by the GA are the locations (x- and y-co-ordinates) of the six cubic spline control points that specify the eye template. An initial random population of chromosomes (each consisting of 6 sets of co-ordinates representing 6 cubic spline control points) is generated. The area corresponding to the extracted iris is removed from the valley map. The sum of the pixels in this modified valley map corresponding in position to the pixels of the eye template, normalised over the length of the eye template, is the fitness function.

The endpoints of the extracted eye contour splines and the centroids of the extracted iris contours represent the eye corners and the pupil centres respectively. The 3D real-world co-ordinates of the eye corner and pupil centre points are obtained using the DLT and the calibration markers. The distances shown in Figure 1 are then calculated. Figure 4 (b) shows extracted eye contours, judged by visual inspection to be representative of the actual features. Figure 4 (c) shows eye contours that underestimate the inner eye corner in images in which sclera and skin cannot be clearly distinguished.



(a) (b) (c)

Figure 4. (a) Eye template; The arrows indicate cubic spline control points. (b) Extracted eye contours: a good fit; (c) Extracted eye contours: underestimation of inner corners.

2.1 Comparison of automatic and manual measurements

Manual selection of points: An investigator selected points by clicking the mouse on corresponding points in each pair of images displayed on a computer monitor. This exercise was carried out 3 times and the average co-ordinates were calculated. The distances shown in Figure 1 were calculated using the three-dimensional real-world co-ordinates of the selected points.

Automatic selection of points: The endpoints of the eye contour splines and the centroids of the irises represent the eye corners and the pupil centres respectively. The distances shown in Figure 1 were calculated using the three-dimensional real-world co-ordinates of the eye corner and pupil centre points.

A comparison of the distances obtained from 46 image pairs using the algorithm for automatic extraction of eye features with those obtained after manual point selection revealed that average differences between the manual and automatic values for PFL and IPD were within 1 mm, while the ICD and OCD showed average differences within 2 mm.

3 Stereo matching

Both the algorithm described above and the manual method of selecting points with a mouse click on a computer monitor, obtain relevant landmark points from each stereo image separately. These points might therefore not be truly homologous, resulting in measurement errors. Feature-based matching of areas of interest in the stereo images to ensure homologous points for measurement would complement both automatic feature extraction and manual point selection. In the former, it could optimise the search space of the GA, while in the latter, it could reduce the time taken to perform the manual task.

A stereo matching technique is used, which matches landmarks selected manually on one image in a stereo pair to the second image [7]. Image contrast is improved and edges obtained using a Canny edge detector are superimposed on the images. The locations of selected landmark points on the right image are copied to the left image. An exhaustive search is performed in a window surrounding each location on the left image to find the pixel whose neighbourhood best matches that of each landmark in the right image. The sum of absolute differences between pixel intensities in 11x11 neighbourhoods centred on the landmark pixel in the right image and pixels in the search window in the left image is used to find the best match. Figure 5 shows the results of the matching technique.

3.1 Comparison of matched and manual points

Landmark coordinates on the left image obtained manually and by matching were compared for 48 image pairs. Manual measurements were obtained as described in section 2.1. Average differences between PFL, IPD, ICD and OCD measurements obtained after stereo matching and those obtained manually were within 1 mm.

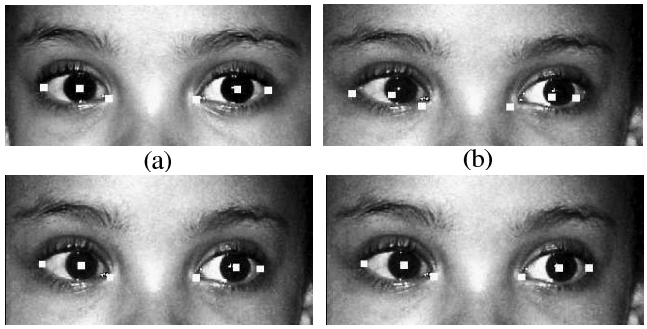


Figure 5. (a) Manually selected landmark points on right image; (b) landmark points copied to left image; (c) manually selected landmark points on left image; (d) matched points on left image.

4 Conclusion

We have developed a stereo-photogrammetric method for obtaining eye measurements for FAS screening, and shown that measurements obtained by selecting landmark points from stereo images by a mouse-click on a computer monitor compare favourably with direct measurements. In an attempt to further reduce the human labour required for large-scale FAS screening, we have presented image processing algorithms for automatic eye feature extraction and feature matching. Average differences between PFLs based on automatically extracted features and those obtained manually were within 1 mm, as were average differences between PFL, IPD, ICD and OCD measurements obtained after stereo matching and those obtained manually.

Since the PFL is the eye feature best suited to the identification of the FAS facial phenotype [2], both the automatic eye feature extraction and stereo matching show promise for the enhancement of large-scale FAS screening. A difference of less than 1 mm between a new and an established method in anthropometry has generally been regarded as acceptable in published work [8]. Although an error of 1 mm could result in a misclassification of a borderline case, such misclassification could be minimized by the use of fuzzy diagnostic borders and the use of a combination of features for diagnosis.

Acknowledgements

This work was conducted with the aid of grants from the NRF (GUN 2053094) and the MRC in South Africa.

References

1. PA May, L Brooke, JP Gossage, et al. "Epidemiology of fetal alcohol syndrome in a South African community in the Western Cape Province", *Am J Public Health* **90**, pp 1905-1912, 2000.
2. SJ Astley, SK Clarren. "A case definition and photographic screening tool for the facial phenotype of fetal alcohol syndrome", *J Pediatrics* **129**, pp 33-41, 1996.
3. DW Smith. "The fetal alcohol syndrome", *Hospital Practice*, October, pp 121-128, 1979.
4. E.M. Meintjes, T.S. Douglas, F. Martinez, et al. "A stereo-photogrammetric method to measure the facial dysmorphology of children in the diagnosis of Fetal Alcohol Syndrome". *Med Eng Phys* **24**, pp 683-689, 2002.
5. Douglas, T. S., Meintjes, E. M., Vaughan, C. L., et al.. "Role of depth in eye distance measurements: Comparison of single and stereo-photogrammetry", *Am J Hum Biol* **15**, pp 573-578, 2003.
6. T.S. Douglas, F. Martinez, E.M. Meintjes, et al. "Eye feature extraction for diagnosing the facial phenotype associated with Fetal Alcohol Syndrome". *Med Biol Eng Comput* **41**, pp 101-106, 2003.
7. V. B. Kostousov, I. L. Molochnikov. "Flexible net approach for stereo matching", *ISPS Proc Photogrammetric Computer Vision* **34(3B)**, pp 126-128, 2002.
8. LG Farkas, CK Deutsch. "Anthropometric determination of craniofacial morphology". *Am J Med Genet* **65**, pp 1-4, 1996.

Automating cell analysis: intracellular particle tracking

Alistair Fitch*, Frank Suess and Patrick Courtney
PerkinElmer Life and Analytical Sciences, Chalfont Road, Seer Green, Beaconsfield, Bucks., HP9 2FX, UK.

Abstract. This paper addresses the quantitative analysis of time series imaging of biological cells. Currently, results are the qualitative visual assessment of video showing the dynamics of the cell. Such results are inherently subjective. Automatic quantitative measurements would allow deeper understanding. In this paper computer vision techniques are applied to obtain quantitative measurements from in-vivo intracellular images.

1 Introduction

This paper is concerned with quantitative analysis of time series images of biological cells. Through recent developments in fluorescent protein labels, a given protein can be highlighted in a live cell [1]. Different proteins can be imaged with different fluorescent labels and selective excitation light and/or emission discrimination. Widefield microscopy collects light from a broad range of z-depths (where z is parallel to the optical axis). Typically widefield microscopes collect light from all z-depths in a cell. Confocal microscopy collects light from a narrow range of z-depths only [2]. Example x, y confocal planes are shown in Figure 1. By shifting the z-depth on a confocal microscope a sample can be directly imaged in all three spatial dimensions. Confocal data is characterised by low signal to noise ratio, noise with a Poisson distribution, high pixel bit depth (12 to 16 bits) and is multidimensional in nature (x, y, z, time, excitation wavelength, emission wavelength).

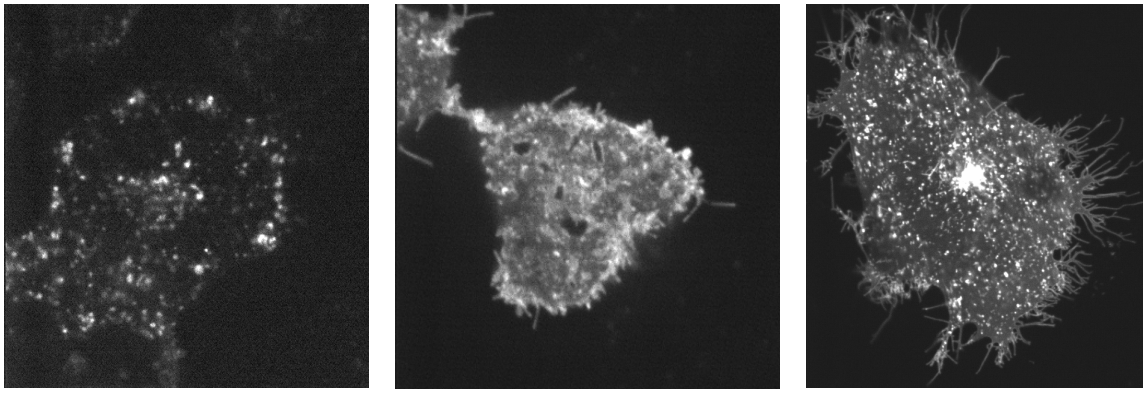


Figure 1. Example confocal z-slices of different cells with fluorescent protein labels.

Recent advances in both fluorescent protein labels and confocal microscopy are allowing a wide range of intracellular dynamics to be studied. Currently, results are the qualitative visual assessment of video showing the dynamics of the cell but such results are inherently subjective. Automatic quantitative measurements would allow deeper understanding. In this paper computer vision techniques are applied to obtain quantitative measurements from in-vivo intracellular images. A system overview is shown in Figure 2.

Use of computer vision techniques to assist quantitative analysis of cellular images covers a wide range of problems and solutions. Here we consider the task of tracking fluorescently labelled proteins. We focus on this task for two reasons. Firstly, quantitative results can be directly measured from the tracks of the fluorescent labels. For example average velocity or total displacement. Secondly, successful tracking of labelled protein is a key building block for further quantitative analysis. For example green fluorescent protein (GFP) labels can highlight one protein and red fluorescent protein (RFP) labels can be used to highlight another protein, tracks of each type of protein allows their interaction to be studied.

The field of computer vision is rich with tracking algorithms. While a full review of these techniques is beyond the scope of this paper, a brief overview of utilised techniques follows: Image processing to enhance required features, pattern recognition to detect features to track, pattern matching to link objects through time, and predict and update cycles in filtering techniques, such as Kalman filtering, to improve tracking [6]. Relevant prior work

* Email for correspondence: alistair.fitch@perkinelmer.com

applied to intracellular imaging includes a comparison of algorithms for single particle tracking [7] and a fuzzy logic based system for multiple particle tracking [8].

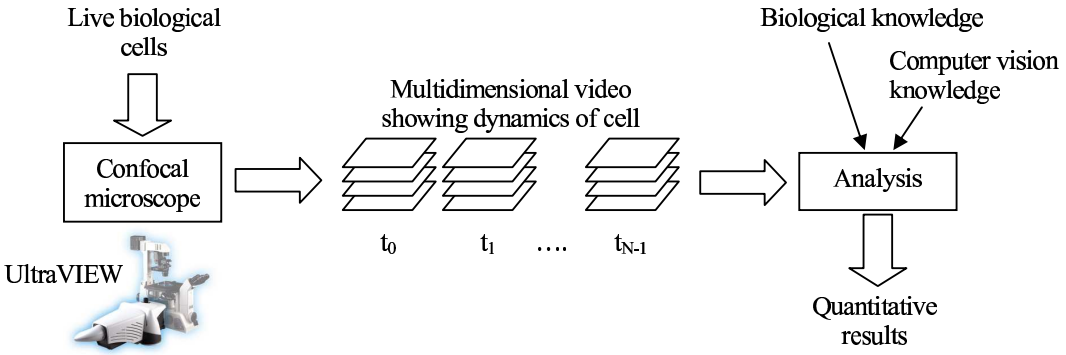


Figure 2. System overview, emphasis is on the use of computer vision for quantitative analysis.

The remainder of the paper is arranged as follows. Our approach is detailed in Section 2. Results from applying the proposed approach to images of cells are shown in Section 3. Examples of the quantitative results obtained are given. Conclusions are drawn in Section 4.

2 Method

We divide the task of two-dimensional intracellular particle tracking into two parts: detection and tracking. Each of these parts are discussed in turn. Significant to the design of both detection and tracking is the simplicity and variation of the target to be tracked. A single fluorescent label is a point light source that will be broadened by the optics of the microscope. Intracellular vesicles contain many fluorescing labels making shape visible. However, the shape of a vesicle may change as it moves. As particles move in z they first become out of focus, then disappear from a confocal plane. Significant intensity variations come from variations in the excitation light source and the photo-bleaching effect. Photo-bleaching is due to damaging of fluorescent labels by the excitation light. Photo-bleaching is observed as an overall reduction in fluorescing intensity over time. Cells may contain thousands of particles.

Detection: The information flow for particle detection is illustrated in Figure 3. A difference filter is used to enhance the particles prior to thresholding. Typically particles cover a small and variable percentage of pixels making automatic thresholding difficult. The user is shown the binary image to assist in setting the threshold. Each contiguous region is considered a candidate particle. The position of a particle is measured as the centroid of the contiguous region. The area and average intensity of each particle candidate is measured. Candidate particles within user specified ranges of area and intensity are promoted to the particle set of given frame. Particle detection is repeated for each frame to produce a set of particles for each frame.

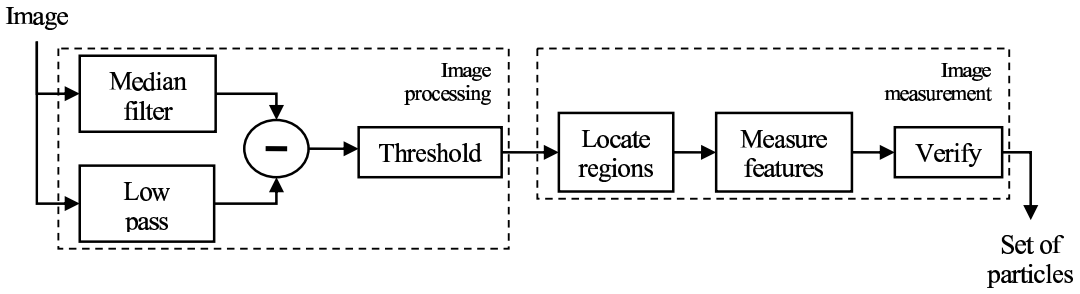


Figure 3. Particle detection information flow.

Tracking: Temporally linking particles to form tracks is based on a powerful matching algorithm that links two sets of particles. The matching algorithm is initially run on temporally adjacent sets of particles, i.e. sets at time *t*

and t+1. To account for missed detections or temporally non-visible particles, the algorithm is then run on the unmatched subset of particles between adjacent but one sets of particles, i.e. sets at time t and t+2. This is iteratively repeated up to a maximum time point gap (Max Frame Difference).

To match sets of particles P_t and P_{t+1} we construct a graph in which each particle is a vertex. It is assumed that particles do not combine. We consider the possibility that each particle in P_t could be matched to each particle in P_{t+1} by constructing a bipartite graph. Each edge in this graph represents a possible linking of a particle in P_t with a particle in P_{t+1}. The weight of each edge is set to the square of the Euclidean distance between the position of the two particles. Our task is now formulated as finding the set of vertices which connects each particle and has a minimum summed weight. This is known as the assignment problem (also referred to as the marriage problem). Given equal size P_t, P_{t+1} and cost function C:$P_t \times P_{t+1} \rightarrow$R the assignment problem is defined as finding the bijection f:$P_t \rightarrow P_{t+1}$ such that (1) is minimised.

$$\sum_{a \in P_t} C(a, f(a)) \tag{1}$$

If P_t and P_{t+1} both have n particles there will be $n!$ possible sets of vertices. Fortunately there is a polynomial time solution - the Hungarian algorithm [3,4]. The case of sets with an unequal number of particles is handled by adding dummy particles. Formulation of tracking as an assignment problem has be applied elsewhere, for example [5]. A key advantage of this approach is that all particles are considered together and a global minimum is found. This is in contrast to iteratively updating individual trackers for each particle

3 Results

Tracking results are shown with data acquired with the UltraVIEW confocal microscope. The x, y, t sequences have been recorded by the University of Bristol. Results of the tracking are shown in Figure 4. Particle tracking is run on the entire image, only a sub-region is displayed here. The full image of frame 38 is shown on the left of Figure 1. Particle detections are displayed as green (mid-grey on greyscale) boxes that enclose the particle detection. The sub-regions in Figure 4 contain a single track which is shown as a white line. Only particle detections for a given frame are displayed on that frame, the complete track is drawn on all frames. Graphs of quantitative measurements taken from this track are shown in Figure 5. From the graph on the left we see that the mean intensity rises and falls along the track, this may indicate the particle moving in and out of the confocal plane. From the graph on the right we see the speed of the particle increasing along the track. Quantitative speed measurements can be used as evidence to classify particles. Note the erratic nature of the velocity, this is typical for tracks.

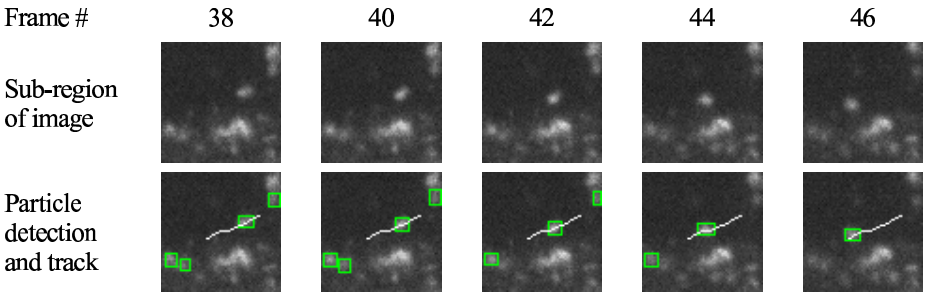


Figure 4. Alternate images of a tracked intracellular particle.

A failure of the tracking algorithm is shown in Figure 6. A full frame from this sequence is shown on the right of Figure 1. The sub-regions in Figure 6 show all tracks in yellow (near white in greyscale), particle detections are shown as for Figure 4. Approximately, the particle of interest starts in the upper right quadrant of the sub-region, moves across to the centre, down to the lower left quadrant, then up to the upper left quadrant. The particle is successfully tracked from frame 5 to frame 78. The track then mistakenly follows a mostly stationary particle in the lower left quadrant. Part of the correct track from lower to upper left quadrants can be seen. Observe the particle detections before and after the mistake. Note in frame 79 the tracked particle is merged with a passing stationary one. Also note that the mistakenly tracked particle is not detected in frame 78 but appears in frame 79. In this situation the solution to the assignment problem mistakenly links the wrong particles.

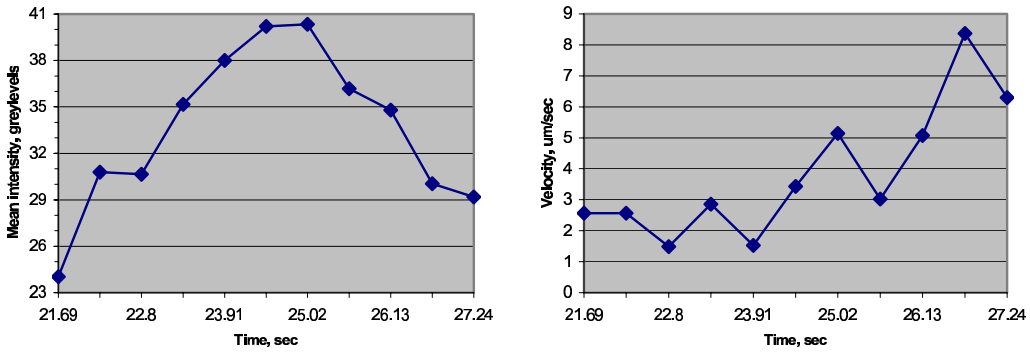


Figure 5. Quantitative measurements from particle track shown in Figure 4.

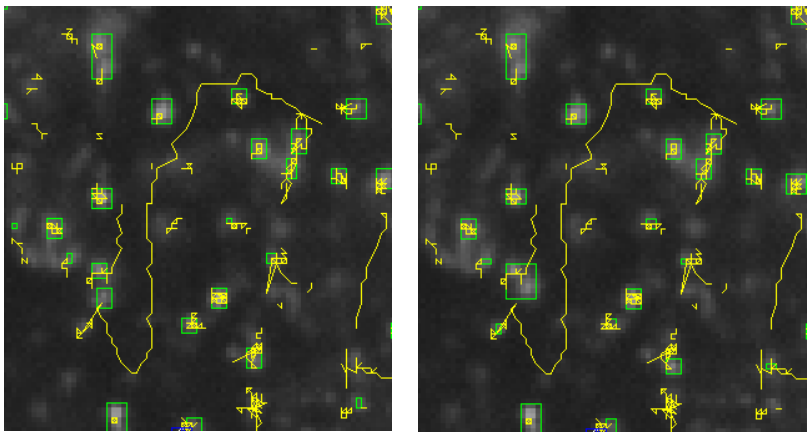


Figure 6. Tracking failure. All tracks shown. Detections shown before and after mistake.

Another open issue is the assessment of intracellular particle tracking algorithms. Of chief concern to users is performance vs. complexity of parameters to setup. This issue is applicable to all tools for quantitative analysis of cells.

4 Conclusion

The general task of automating cell analysis is presented. The paper focuses on the specific requirement of intracellular particle tracking. An implementation of an automatic intracellular particle tracker is described. Quantitative results are presented.

Acknowledgments

Our thanks to Peter Cullen and his lab at University of Bristol for providing the images shown in this paper.

References

1. A. Miyawaki, A. Sawano and T. Kogure, "Lighting up cells: labelling proteins with fluorophores", Imaging in Cell Biology, Supplement to Nature Cell Biology, vol. 5, pp S1-S7, 2003
2. D.J. Stephens & V.J. Allan. "Light microscopy techniques for live cell imaging", Science, 300:82-86, 2003
3. H.W. Kuhn, "The Hungarian method for the assignment problem." Naval Research Logistics Quarterly, 2:83-97, 1955
4. D.E. Knuth, "The Stanford GraphBase: A platform for computational computing, Addison-Wesley, 1993
5. C.J. Veenman, M.J.T. Reinders and E. Backer, "Resolving motion correspondences for densely moving points", IEEE Transactions on Pattern Analysis and Machine Intelligence, 23:54-72, 2001.
6. M. Sonka, V. Hlavac and R. Boyle, "Image Processing: Analysis and Machine Vision", Brooks Cole, 1998
7. M.K. Cheezum, W.F. Walker and W.H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles", Biophysical Journal, 81:2378-2388, 2001.
8. W. Tvaruskó, M. Bentele, T. Misteli, R. Rudolf, C.Kaether, D.L. Spector, H.H. Gerdes and R. Eils, "Time-resolved analysis and visualization of dynamic processes in living cells", Proc. Natl. Acad. Sci. USA, 96:7950-7955, 1999

Image-based identification of hibernating myocardium

Nicholas M.I. Noble[a], Derek L.G. Hill[a]*, Marcel Breeuwer[b] and Reza Razavi[a]

[a]Computer Imaging Science Group, Guy's Hospital, Kings College London, UK,
[b]Philips Medical Systems, Medical Imaging Information Technology - Advanced Development

Abstract. Delayed enhancement imaging is a recently described technique that enables for the first time, the direct observation of areas of myocardium that have scarred following infarction. When this information is combined with information about myocardial contraction, areas that are neither dead, nor contracting can be identified. Such areas will resume contraction following revascularisation (hibernating myocardium). The identification of such areas is consequently of great interest to clinicians. This paper describes how registration can be used to align the images prior to the identification of areas that will benefit from revascularisation. Patient data is used to demonstrate image alignment and image-derived information combination. This is then mapped onto patient-specific 2D and 3D representations of the heart.

1 Introduction

It has recently been observed that gadolinium-based contrast agents have substantially different uptake characteristics in regions of scarred myocardium as opposed to in living areas [1]. Whereas these contrast agents wash into and out of the vascular system of living myocardium in a few heart beats, in areas of scarring, dead cardiomyocytes are replaced by a matrix of collagenous fibres —scarred myocardium— into which contrast agent accumulates 15-20 minutes post-injection. Called delayed enhancement imaging, this technique enables for the first time, the direct non-invasive observation of scarred myocardium. When information about areas of scarred myocardium is combined with knowledge of local myocardial contraction or perfusion, areas of myocardium that are neither dead nor contracting —*hibernating myocardium*— can be distinguished. Hibernating myocardium has the potential to resume contracting if revascularised [2]. Its location is thus of considerable importance to clinicians planning revascularisation.

Currently, to identify areas of myocardial scar, clinicians use either Positron Emission Tomography (PET), or Single Photon Emission Computed Tomography (SPECT). Scar size measurements using delayed enhancement imaging have been shown to correlate closely with those measured using PET [3]. In addition MR has been shown to systematically detect small sub-endocardial infarcts that were not identified using SPECT [4]. The higher voxel resolution and non-ionising nature of MR make it well suited for the identification of hibernating myocardium. An emergent technology that is competing with delayed enhancement is contrast enhanced echocardiography. It remains to be seen which technique emerges as the clinical method of choice.

The identification of hibernating myocardium from MR images can be performed in two ways. The first entails finding areas that demonstrate both perfusion defects in first-pass perfusion images, and an absence of scar in delayed enhancement images. Recently, Breeuwer *et al.* investigated aligning such images [5]. The second approach involves finding areas of myocardium that exhibit reduced contraction in cine anatomical images, and are not scarred in the delayed enhancement image. To identify such regions the images need to be aligned, this is normally performed mentally by the observer. Differences in the position of the heart due to inter-image motion and inconsistent breath-hold positions confound this already difficult task. This paper will investigate the alignment of cine anatomical images with delayed enhancement images, before identifying areas of hibernating myocardium.

2 Data

Short axis ECG triggered steady state free precession images with SENSE factor 2 were obtained in 2 patients undergoing cardiac MRI for the investigation of coronary artery disease. Eight contiguous slices, taken in up to three breath-holds, were imaged with: slice thickness 8 - 10 mm, field of view 350×344 - 390×390 mm, acquisition matrix 192×192 with 120% phase encode direction sampling, reconstructed to 256×256 giving a resolution of 1.8 mm×1.8 mm – 2×2 mm, with 20 phases in the cardiac cycle, flip angle 50 - 55 °, TE 1.56 - 1.68 ms and TR 3.11 - 3.37 ms. A 0.4 mmol/kg body weight bolus of gadolinium DPTA was then administered intravenously. Fifteen minutes post-injection, a single end diastolic image was acquired. The images were acquired on a Philips Gyroscan Intera 1.5 T with master gradients, using a 5 element cardiac synergy coil and vector ECG.

3 Methods

Before information about contraction and scarring can be combined, misalignment of the heart between the cine anatomical and delayed enhancement images, due to inter-scan motion and breath-hold inconsistencies, must be compensated for. Both the delayed enhancement and cine anatomical images are acquired using ECG gating, the delayed enhancement image should consequently be at the same point in the cardiac cycle as the end diastolic cine

*Corresponding author D.L.G. Hill: Derek.Hill@kcl.ac.uk

anatomical image. A rigid registration should hence suffice to align the heart, however, rigid registration of the entire image will only recover differences in bodily position and not those attributable to inconsistent diaphragm position. The registration of a Region Of Interest (ROI) in the delayed enhancement image —defined by manually segmenting the epicardial surface— overcomes this problem. The end diastolic cine anatomical image was thus rigidly registered to the ROI of the delayed enhancement image. Normalised mutual information [6] was chosen as the similarity measure because of its ability to align areas that include contrast in one image but not the other. The transformation produced by the registration may well result in some degree of through-plane motion. Because of the massive voxel anisotropy inherent in these images, a smooth 3D contour must hence be created prior to its transformation. Shape-based interpolation [7] of binary volumes produced by segmentation of anatomical end diastolic epicardial and endocardial surfaces provides approximately isotropic voxels. Gaussian blurring ($\sigma = 1$mm) of the interpolated binary volume followed by Marching Cubes [8], then extracted a smooth 3D surface. Figure 1 shows sections through these contours overlaid on the delayed enhancement image both prior to registration and following registration and transformation. To enable the later assessment of wall thickening, end systolic epicardial and endocardial segmentations were also prepared as described above and transformed by the same transformation. The end systolic image should be correctly aligned with the end diastolic image because it is acquired with the same breath-holds, there is thus no need for alignment of the end diastolic images to the end systolic cine images.

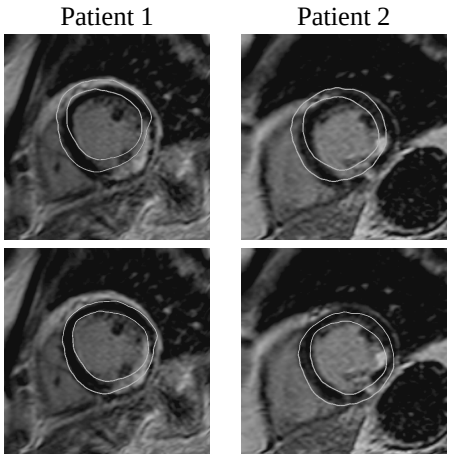


Figure 1. Manually created end diastolic anatomical contours overlaid onto the delayed enhancement images prior to registration (*top row*), and following registration (*bottom row*).

A bullseye plot is a series of contiguous concentric rings, segments of each ring are colour coded according to the parameter calculated for corresponding segments of myocardium. Each of the concentric rings in the plot represents information originating from a particular slice, the slice nearest the apex representing the innermost ring, and the outermost ring information from the slice nearest the base. The in-plane angular location of the information determines the angular location at which it is positioned in the bullseye plot. In this case, parameters for each segment of the bullseye plot were measured in-slice at equi-spaced angular locations about the in-slice centroid of the delayed enhancement endocardial surface.

Areas of scar were identified manually and the percentage transmurality of the scar plotted on a bullseye —Fig. 2, second column. The percentage wall thickening (a metric routinely employed in echocardiography) was calculated from the transformed end diastolic and end systolic contours, and mapped onto bullseyes —Fig. 2, second column. Because the contours have been aligned prior to being mapped onto the bullseye plots, for the purposes of data combination, it will be sufficient to combine the information associated with the segments of the bullseye plots.

To combine the information from the two bullseye plots, a function of the wall thickening (x) and transmurality (y) needs to be defined. This function should be maximum in areas of low thickening and low scarring, indicating the likely location of hibernating myocardium —candidate areas for revascularisation. The function should in addition be minimum in areas of maximal thickening and also in areas of maximal transmurality —areas, the revascularisation of which, will not benefit the patient. One such candidacy function can be described by $z = (1 - x)(1 - y)$. This function was used to combine the wall thickening and transmurality information on bullseye plots for each patient —Fig. 2, third column. It should be noted that although this function has the desired features as described above, its behaviour away from these points is a somewhat arbitrary estimation of the benefit associated with the revascularisation of such areas. The true revascularisation candidacy function, can only be determined experimentally by the assessment of both pre- and post-revascularisation images.

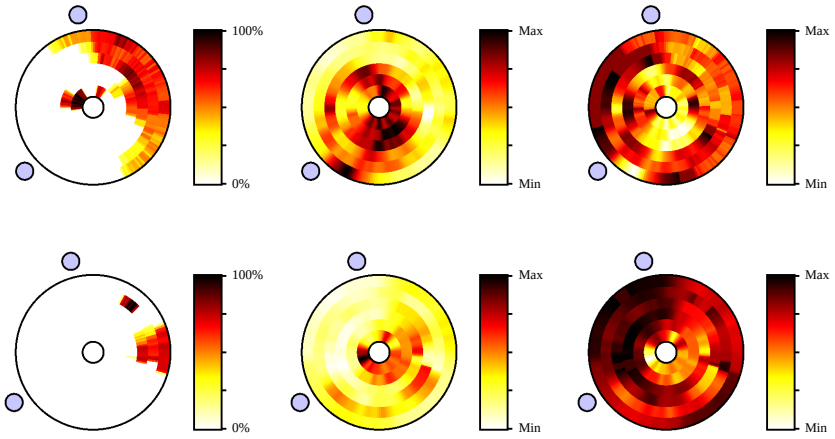


Figure 2. Bullseye plots showing; transmurality *first column*), wall thickening (*second column*), and candidate areas for revascularisation (*third column*), for patients 1 and 2 (*top and bottom rows*). Dots indicate the location of the conjunction of the left and right ventricles.

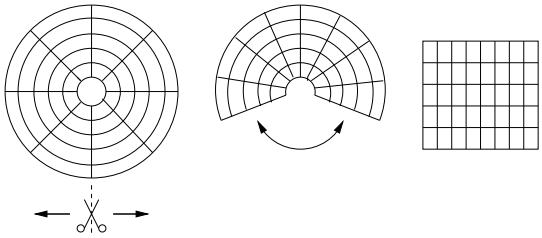


Figure 3. The unwrapping of the bullseye plot into a rectangular image.

Although the bullseye plot permits two-dimensional observation of the location of three-dimensional data, it is somewhat abstract. An easier to interpret representation would be if the combined information were visualised on a 3D representation of the myocardium. To thus represent the combined wall thickening and transmurality information, a coordinate system that can be used to translate between the bullseye plot and the surface must be defined. Segments in the bullseye plot can be easily identified using a polar representation, on the surface this corresponds to a (z, θ) coordinate system —as opposed to an (r, θ) system on the bullseye plot. The location of the centroid of any of the 3D surface's facets on the bullseye can be identified from its z-position and its theta value calculated with respect to its in-slice centroid. This is a continuous coordinate system and the resulting location in the bullseye plot may consequently be at a non-integer position. To enable interpolation of intermediate values from the bullseye plot, it was unwrapped (see Fig. 3). Cubic b-spline interpolation —wrapped around at the image edges to account for radial continuity— then enabled smooth $C2$ interpolation of the bullseye plot. Examples of the bullseye plots for transmurality, wall thickening and the combined function are shown mapped onto prepared segmentations of the delayed enhancement epicardial surface in Fig. 4.

4 Results

The contour overlays in Fig. 1 clearly demonstrate that for patients 1 and 2, the registration has aligned the hearts very well. For patient 3 however, there is a substantial registration error. It would seem that the registration algorithm has aligned the enhanced region with the blood pool in the cine anatomical image.

Figure 2 shows bullseye plots of transmurality, wall thickening and candidate areas for revascularisation for all three patients. It can be seen that patient 1 has a sizable amount of scar tissue in the basal and mid anterior segments of the heart, a corresponding area of reduced thickening can be seen in the wall thickening plot. The combined plot does not indicate any substantial areas with high revascularisation candidacy, although a small region may be seen in the basal septal areas. This is nicely illustrated in Fig. 4, where for the first two positions, one can clearly see the reduced wall thickening in areas of scar. Patient 2 has two small areas of scar in the basal anterior segments, the wall thickening bullseye plot shows reduced thickening in this area and also in the septal segments. When combined, it can be seen that there is a large candidate area for revascularisation in the basal and mid septal segments. This is also seen in Fig. 4.

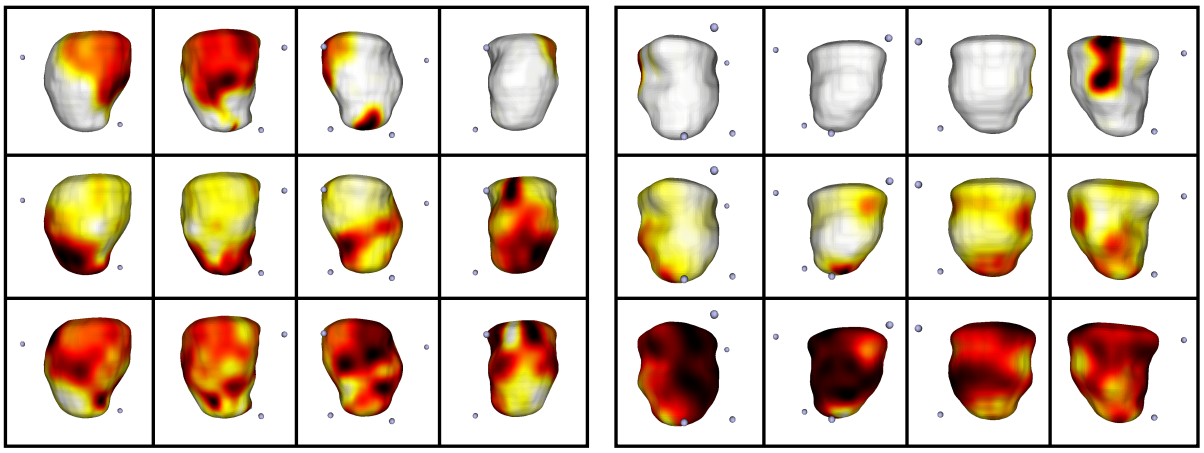


Figure 4. Mappings of transmurality (*top row*), wall thickening (*middle row*) and candidate areas for revascularisation (*bottom row*) for patients 1 (*left*) and 2 (*right*). Spheres indicate the location of the conjunction of the left and right ventricles. Colour scales are the same as employed in figure 2.

5 Discussion and Conclusion

This paper has addressed the task of automatically combining information from cine anatomical and delayed enhancement images and, to the author's knowledge, is the first paper to attempt this. Rigid registration using normalised mutual information of a region of interest was shown be to sufficient to compensate for both inter-scan patient movement and inconsistent breath-hold positions. Following registration, the location of manually identified areas of scarring was then combined with manually determined wall thickening information using a combination function. Although at this stage the combination function is rather arbitrary, it enables candidate areas for revascularisation to be highlighted for the first time. The results of the information combination were shown on both 2D bullseye representations of the heart, and on 3D representations of the myocardial surface. Future work required to make this technique clinically employable would be the empirical identification of a revascularisation candidacy function, this will require a substantial set of both pre- and post-revascularisation images. Following this, the technique could be validated by comparison with results derived from PET—the current gold standard.

Acknowledgements

We are grateful to Philips Medical Systems Nederland B.V. Medical Imaging Information Technology - Advanced Development for funding this work, and to Professor Hawkes and all those at the Computer Imaging Sciences Group for their assistance.

References

1. R.J. Kim, D.S. Fieno, T.B. Parrish, K. Harris, E. Chen, O. Simonetti, Bundy J., J.P. Finn, F.J. Klocke, and R.M. Judd. Relationship of MRI Delayed Contrast Enhancement to Irreversible Injury, Infarct age, and Contractile Function. *Circulation*, 100:1992–2002, 1999.
2. S. Rahimtoola. The Hibernating Myocardium. *Am. Heart J.*, 117(1):211–230, 1989.
3. C. Klein, S.G. Nekolla, F.M. Bengel, M. Momose, A. Sammer, F. Haas, B. Schnackenburg, W. Wolfram Delius, H. Mudra, D. Wolfram, and M. Schwaiger. Assessment of myocardial viability with contrast enhanced magnetic resonance imaging: comparison with positron emission tomography. *Circulation*, 105(2):162–167, 2002.
4. A. Wagner, H. Mahrholdt, T.A. Holly, M.D. Elliott, M. Regenfus, M. Parker, F. Klocke, R.O. Bonow, R.J. Kim, and R.M. Judd. Contrast-enhanced mri and routine single photon emission computed tomography (SPECT) perfusion imaging for detection of subendocardial myocardial infarcts: an imaging study. *Lancet*, 361:374–379, 2003.
5. M. Breeuwer, R. Muthupillai, S. Flamm, E. Nagel, I. Paetsch, J. Ridgeway, and S. Plein. Combining first-pass and late-enhancement myocardial perfusion analysis. In *Proc. Society of Cardiovascular Magentic Resonance*, 2003.
6. C. Studholme, D.L.G. Hill, and D.J. Hawkes. An Overlap Invariant Entropy Measure of 3D Medical Image Alignment. *Pattern Recognition*, 32:71–86, 1999.
7. G.T. Herman, J. Zheng, and C.A. Bucholtz. Shape-Based Interpolation. *IEEE Comput. Graph. Appl.*, pages 69–79, 1992.
8. W.E. Lorensen and H.E. Cline. Marching Cubes: A High Resolution 3D Surface Reconstruction Algorithm. *Comput. Graph.*, 21(4):163–169, 1987.
9. V. Positano, A. Pingitore, M.F. Santarelli, M. Lombardi, L. Landini, and A. Benassi. Quantitative 3D assessment of myocardial viability with MRI delayed contrast enhancement. In *Computers in Cardiology*, Thessalonika, Greece, 2003.

Creation of patient-specific CFD models by morphing a previously-meshed reference geometry using image registration

David M Jones[a *], D Rodney Hose[a], Patricia V Lawford[a], Derek L G Hill[b], Reza S Razavi[b] and David C Barber[a]

[a]Medical Physics, University of Sheffield, [b]Division of Imaging Sciences, King's College London

1 Introduction

The creation of realistic computational fluid dynamic (CFD) models from medical images has been the subject of much research [1]. However, if simulations are to be performed on a routine basis in a clinical setting, the user interaction required to create the CFD model must be reduced by automating the mesh generation process as far as possible, and this is the subject of this paper. Furthermore, high performance computing required for transient CFD simulation must be accessible, possibly using Grid technology [2]. We describe a technique for rapidly generating patient-specific meshes, and the potential to validate these using flow and invasive pressure measurements made during catheterisation under MR guidance.

2 Automatic mesh generation

Typically, the stages in generating a CFD mesh – segmentation, profile extraction, surface generation, and meshing - are performed for every new CFD model that is created, often with significant user intervention. We propose a novel technique for creating patient-specific models, in which the time consuming step of creating a CFD mesh for a given arterial topology need only be performed once for a 'reference patient'. This mesh is then morphed to any individual by image registration [3] of the reference CT or MR scan data to the patient scan.

2.1 Creation of Mesh for Reference Patient

An idealised representation of the real geometry is defined in terms of geometric primitives, such as cylinders, and then meshed. From the mesh, a pseudo-binary image is created by identifying which points on a regular 3D grid lie inside the mesh surface (intensity 1), and which lie outside (intensity 0). The scan data for the reference patient is then segmented using any appropriate technique to produce a second binary image. Because this process need only be performed once, we performed the segmentation manually. The two binary images were blurred using a Gaussian filter with a FWHM of 4 voxels, and then registered to obtain a mapping from the idealised geometry to the real geometry. By applying the registration mapping to the idealised mesh, we obtain a mesh for the reference patient. This process is illustrated in Figure 1. This technique generates an appropriate distribution of mesh density and the inlets and outlets can be automatically identified by reference to the original geometric primitives.

2.2 Creation of Patient-Specific Meshes

Subsequent scans (of the same patient at different times, or different patients) can be registered directly (at image level) to the original reference scan, and the mapping thus obtained applied to the reference mesh to obtain a scan-specific mesh. Processes such as wall movement or dilation in a dynamic scan sequence are obtained naturally and directly from the registration mapping. The segmentation of each image except for the reference one is performed effectively and automatically by non-linear image registration.

[*] *Corresponding author*: Room I100, Royal Hallamshire Hospital, Glossop Road, Sheffield, S10 2JF, UK
E-mail address: d.jones@sheffield.ac.uk

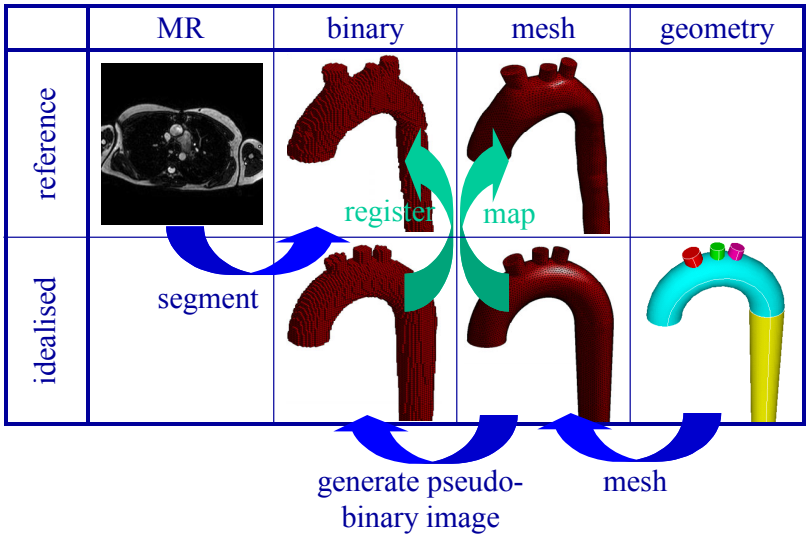
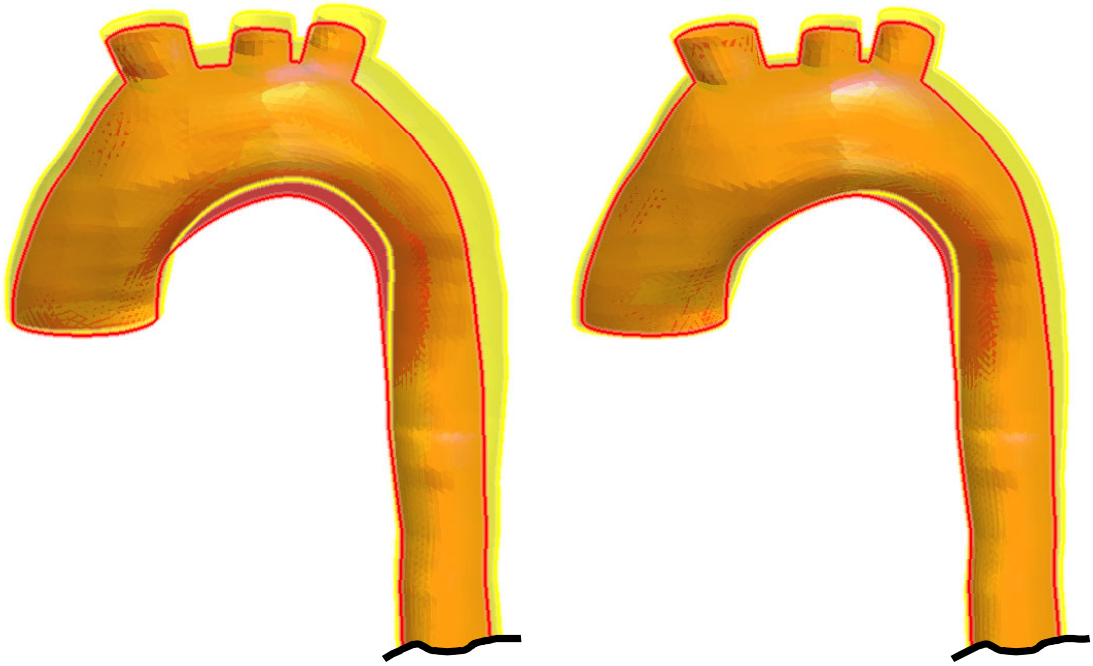


Figure 1. Steps involved in generating the reference mesh

3 Example Application

We have applied the technique to the thoracic aorta of a patient with congenital narrowing of the aorta (coarctation), who underwent dynamic MR imaging, flow imaging, and invasive pressure measurements in the integrated MR and x-ray (XMR) catheterisation lab at Guy's Hospital. The thoracic aorta is a good exemplar blood vessel because it is a complex shape and has significant motion, and is large enough that the lumen can be reliably imaged using MRI. The ability to generate patient-specific (rather than generic) models is particularly important here because of the high variability in anatomy in patients with congenital heart disease. Invasive investigation in patients with mild coarctation is not normally justified. Therefore management of these patients would be improved if changes in pressure and velocity distribution in the aorta with exercise could be predicted by MR-based CFD and structural simulation alone.

A dynamic series of seven MR scans was acquired using three-dimensional steady-state free precession volume scanning (TR = 3.2ms, TE = 1.6ms, multi-breath hold, multi-phase technique, with a reconstructed voxel size of 1.25mm × 1.25mm × 1mm). The first scan in the dynamic sequence is treated as the 'reference' image, and scans at subsequent time points are registered to this one. Figure 2 shows the meshing technique applied to the dynamic sequence from this patient.



(a) red: reference at 54ms, yellow: 264ms (b) red: reference at 54ms, yellow: 474ms

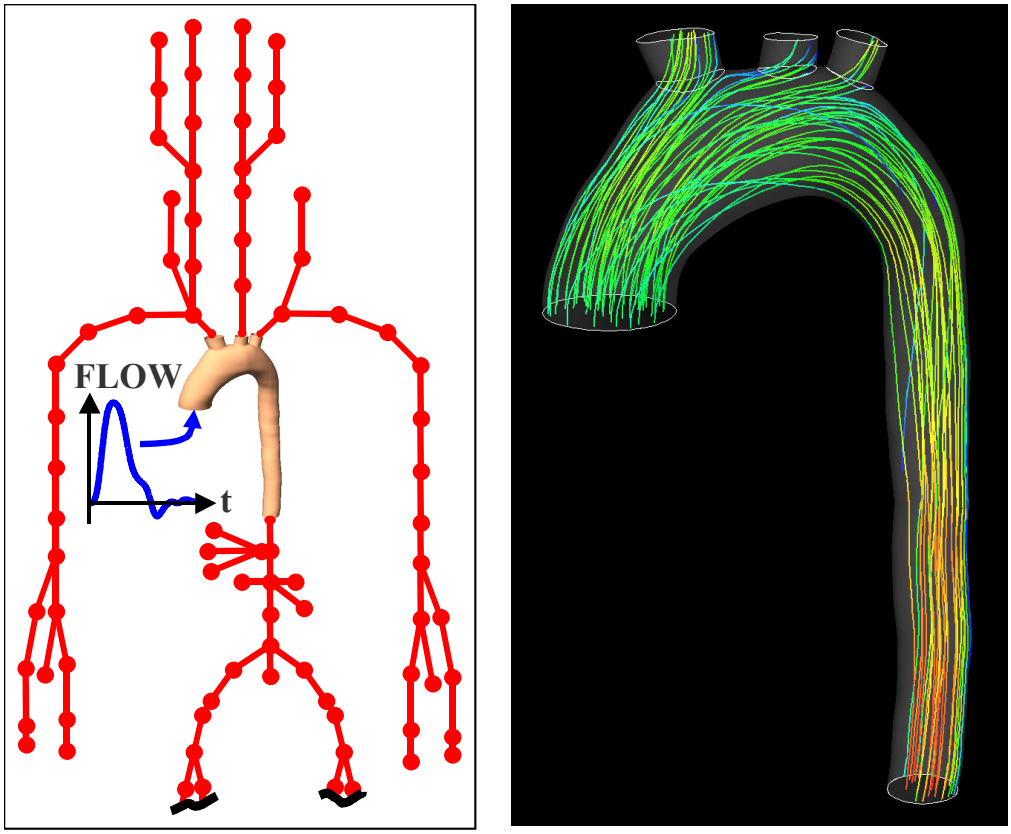
Figure 2. Comparisons of reference model (red) with model for a later time point (yellow). Times after ECG trigger are given.

4 Computational Fluid Dynamics Modelling

Geometrical information obtained from the morphed mesh may in itself be important – for example it may be used to determine the amount of aortic narrowing in patients with coarctation. However, the mesh can also be used in CFD simulations to determine the pattern of blood flow in the arterial segment. In this case, the inlet flow to the aorta model was specified using data from a flow-sensitive Fast-Field MR Echo sequence acquired in the same session as the anatomical scan. The outlets of the model were coupled to compartmental models that represent the downstream circulatory system. For the compartmental models, we used a digital implementation of sections of the model proposed by Westerhof et. al. [4], which is an electronic analogue of the human systemic arterial tree. It consists of 121 arterial segments, with each segment consisting of a number of linear passive electrical components. CFD simulation of the 3D aorta model is performed using CFX (ANSYS Inc.). Because of the time-varying nature of blood flow, a transient solution, typically with timesteps of 0.01s or lower, must be performed, and this is very computationally demanding. If blood flow simulations are to be performed routinely in a clinical setting as part of patient management, then high performance computing must be made more accessible. This could be achieved through the use of Grid technology, and the combination of CFX coupled to compartment systems has implemented on the GEMSS testbed Grid [2].

For the current model, converged solutions have been obtained for three-quarters of the period of systole – divergence at the end of the deceleration phase is under investigation. Figure 3 shows the reference mesh coupled to the compartment models and solution streamlines at peak systolic flow. The solution of the compartment models allows the pressure and flow at downstream locations in the circulatory system to be determined. Also, the compartment models provide better boundary conditions for the 3D model – if the compartments are replaced by simple zero-pressure boundaries, then the flow for the three arteries leaving the aortic arch is too high.

In the future we intend to develop these models to allow the time-varying wall motion to be imposed on the CFX model, or, alternatively, be determined by coupling the wall boundary to a structural finite element model representing the elastic properties of the vessel wall, as described by Hose et. al.[5].



(a) 3D mesh coupled to downstream compartments

(b) Solution at peak systolic flow

Figure 3. Computational Fluid Dynamics Simulation

References

1. D.A. Steinman "Image-Based Computational Fluid Dynamics Modeling in Realistic Arterial Geometries", *Annals of Biomedical Engineering* **30,** pp 483-497, 2002.
2. D.M. Jones, J.W. Fenner, G. Berti, et. al. "The GEMSS Grid: An evolving HPC environment for medical applications". In *Proc. HealthGrid*, Clermont Ferrand 2004.
3. S. Mehta, D.C. Barber, E. van Beek, et. al. "Delineation of the prostate capsule in 3D-Trans Rectal Ultrasound images using image registration". In *Proceedings of Medical Image Understanding and Analysis*, pp. 29-32, BMVA, 2003.
4. N. Westerhof, F. Bosman, C.J. De Vries, et. al. "Analog studies of the human systemic arterial tree", *Journal of Biomechanics* **2**, pp. 121-143, 1969.
5. D.R. Hose, P.V. Lawford, A.J. Narracott et. al. "Fluid-solid interaction: benchmarking of an external coupling of ANSYS with CFX for cardiovascular applications", *Journal of Medical Engineering & Technology*, **27** pp. 23-31, 2003

Modelling Normality and Functionality of the Lungs

Harbir Singh [a*], Michael Crawford[b], John Curtin[b] and Reyer Zwiggelaar[c]

[a]School of Computing Sciences, University of East Anglia, Norwich, England.
[b]Norfolk and Norwich University Hospital, Norwich, England.
[c] Department of Computer Science, University of Wales, Aberystwyth, England.

Abstract. The high prevalence and mortality of lung cancer has encouraged studies in the early detection of lung cancer. Driven by the need for a better understanding of the anatomical functionality and normality of the lung using CT data, we highlight our findings related to application areas like lung segmentation, density estimation and segmentation of anatomical features of the lung such as the airway tree structure.

1 Introduction

Accurate segmentation of the lungs into physiological structures such as the intra-thoracic airway tree and the pulmonary structure is a crucial part of lung automated diagnosis system. As part of modelling the normality and functionality of the lungs, in this paper we concentrate on lung segmentation, density estimation and lung airway tree segmentation. The next sections describe application areas such as the effect of patient orientation on lung density and the task of airway tree segmentation.

2 Effect of Patient Orientation on Lung Density Estimation

Lung CT data can be obtained with the patient in the prone or supine position. To quantify the effect patient orientation (prone or supine) has on lung density, the left and right lungs were segmented with respect to the background in the CT lung images. To avoid inter- and intra-observer variation, an automatic approach was developed which incorporates a number of distinct steps.

The standard lung window (center and width respectively -500 and 1500 Hu) was used as the basis of the segmentation process (see Fig. 1a). The image was subsequently thresholded using simple histogram based peak separation. The next stage of the segmentation process involved a morphological closing. The structuring element was a 3×3 square. The main aim of this step was to remove noise aspects and small anatomical structures. After the closing operation connected components were labelled. The two largest regions were selected as being the two lung areas (see Fig. 1b).



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

Figure 1. Automated Lung Segmentation: (a) example CT slice, (b) segmented left and right lung areas, (c) masked left lung, and (d) masked right lung.

The case of joint lungs, where the anterior and the central junctions of the CT slices had weak contrast making it difficult to separate the left and right lungs was handled using iterative erosion until two major lung regions were obtained. The two lung regions where subsequently dilated, restricted by the original segmentation, to obtain the original lung area [1].

The output of the segmentation stage is the left and right lung masks. The correlation between manual and automatic segmentation can be found in Fig. 2. The masks in combination with the original DICOM data result in left and right lung regions (see Fig. 1c,d). It can be seen that the left and right lung regions contain dense vessels and airways which appear as areas of white blobs and dark patches surrounded by white vessel walls. These do not form a part of the lung tissue needed for the density estimation and were removed by thresholding at a level of -100Hu. The mean pixel value of each lung region was calculated and was taken as an estimate of lung density.

*Email: harbir.singh@uea.ac.uk, michael.crawford@nnuh.nhs.uk, john.curtin@nnuh.nhs.uk, rrz@aber.ac.uk

To evaluate the effect of patient orientation, each lung area was divided into three portions, namely anterior, central and posterior where each portion was one third the height of the lung. The densities of the regions were calculated and correlated to identify any pattern of density variation which could be attributed to the orientation. We also compared the supine and prone densities of the full right and left lungs.

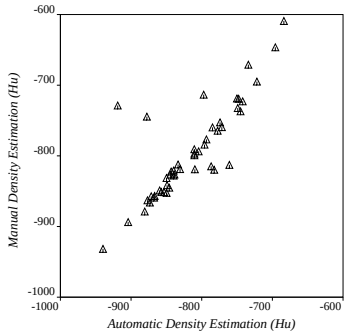


Figure 2. Lung segmentation: manual versus automatic.

2.1 Results and Discussion

The data comprised supine and prone inspiration and expiration CT scans of fifteen patients obtained from the Norfolk and Norwich University Hospital. Scans were obtained in the supine position at 1 cm intervals during inspiration and at 3 levels during expiration, and in the prone position at 3 cm intervals during inspiration and at 3 levels during expiration. From each patient twelve scans were obtained; three supine inspiration, three supine expiration, three prone inspiration and three prone expiration. Each set of the scans was acquired at the upper, middle and lower portions of the chest region. The inspiration/expiration and prone/supine levels were matched as best as possible. All image values in this paper are represented in Hounsfield units (Hu).

An overview of the results can be found in Fig. 3. For both the left and right lung a low correlation coefficient (respectively 0.631 and 0.619) was obtained. In both cases the supine density estimate tends to be lower when compared to the prone density estimate. These initial results indicate that patient orientation does matter when considering density estimation. However, at the same time it should be noted that the resulting correlation between prone and supine estimates might be high enough from a clinical point of view (this needs further investigation). It is expected that the non-uniform deformation of the lung lobes when comparing the supine and prone position of the patient needs to be taken into account.

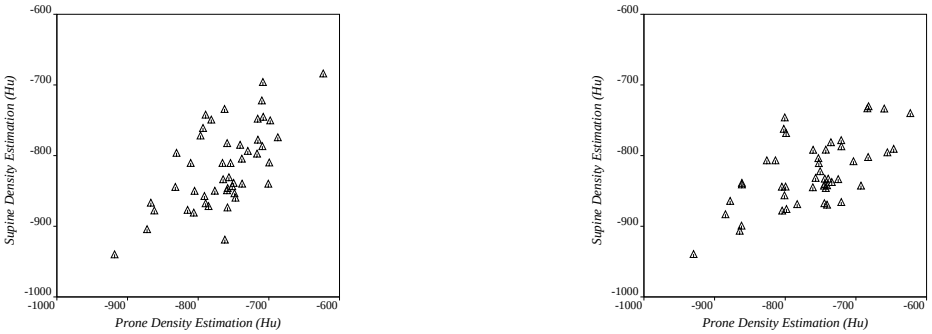


Figure 3. Automatic density estimation (a) left lung, supine versus prone and (b) right lung, supine versus prone.

The non-uniform deformation aspects are further emphasized in Tab. 1. This table summarizes the correlation between the anterior, central and posterior regions of the lungs for supine and prone inspiration/expiration. As expected, there is a larger variation for the inspiration comparison, where it is expected that the non-linear deformation is more pronounced. A simple explanation for the lack of correlation with respect to the orientation of the patient could be the mismatch between slices from the supine and prone positions. The prone inspiration data consisted of slices at 3 cm intervals. For the supine inspiration data the slice distance was 1 cm. From these inspiration slices three were manually selected to match three expiration slices. There is no certainty that the prone and supine slices contain identical anatomical information. A more robust approach would be based on estimating a three dimensional density distribution from the inspiration data and match the supine and prone data and subsequently

match the three expiration slices to these three dimensional density distributions.

Table 1. Regional correlation coefficients. S: supine, I: inspiration, P: prone, and E: expiration

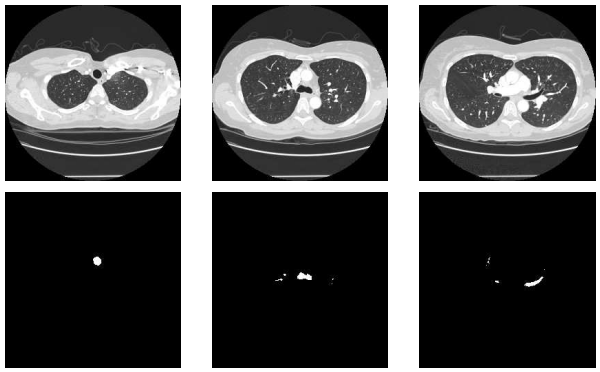
Scans	Correlation Coefficient		
	Anterior	Central	Posterior
SI-PI	0.519	0.605	0.637
SE-PE	0.592	0.595	0.610

3 Airway Tree Segmentation

Providing a non-invasive tool for 3D reconstruction of anatomical structures such as the bronchial tree from 2D and 3D data is a challenging task for computer vision in medical imaging [2]. The airway tree is a pipe structure starting from the end of the trachea and branching like a tree into the lungs. Often intensity inhomogeneity and partial volume effects make it difficult to identify small and finer details at the lower end of the tree. Once the airway tree has been extracted, quantitative analysis can be performed to evaluate tree structure and function which is important for examining physiologic and pathological conditions such as stenosis and tumors. In this section we present and discuss the results obtained using a gain-based 3D region growing approach for segmenting the airway tree. For a detailed discussion on the methodology we refer to [3].

3.1 Results and Discussion

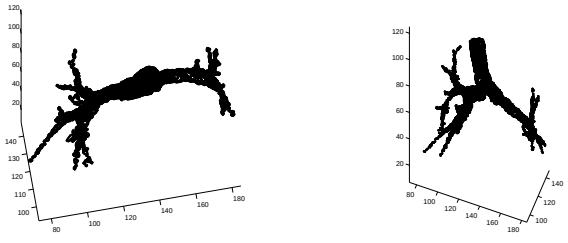
Using volumetric HRCT lung data comprising 125 transversal scans, 256x256 in dimension, Fig. 4 shows the processing results using 2D slices representing the top, middle and lower section of the bronchial tree. The 3D algorithm successfully identifies the bronchus area. Fig. 4a shows part of the trachea, Fig. 4b shows the trachea before splitting and Fig. 4c shows the airways when the trachea has split.



(a) (b) (c)

Figure 4. 2D processing results showing original images and extracted bronchus area.

Fig. 5 shows the 3D visualization of a typical segmented airway tree from the top and side angle. The algorithm has been successful in identifying the branching and the structure of the airway tree to a high degree of accuracy.



(a) (b)

Figure 5. Visualization of airway trees at T=53, G=50%, displaying (a) top view, (b) side view.

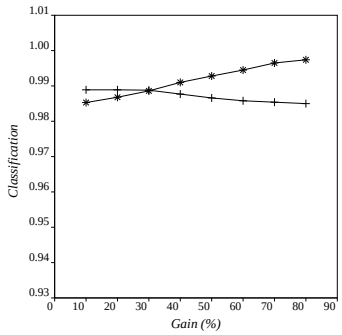


Figure 6. Tree segmentation: evaluating performance of heuristic gain measure. Sensitivity(+) and specificity(*) as a function of gain for threshold T=57.

As part of the evaluation, the automated results were compared to the manually segmented tree and the effect of gain measure was quantified (see Fig. 6). The current evaluation methodology has a bias towards the initial levels in the airway tree as these contain relatively more pixels than the final levels of the tree. A level dependant evaluation is part of our future research directions. It should also be mentioned that the gain-based algorithm is not fail proof as it encounters the problem of leakage which affects all segmentations of the airway tree based on the region growing approach. Even though the heuristic gain was used in this case, it was not able to prevent leakage of the grow process into the lung parenchyma. The heuristic gain at best introduces an additional control facility in the extraction process and combined with other parameters such as explosion of volume and other features of the regions of interest [4], can help in achieving better segmentation results.

Currently, our research concentrates on the development of a modified approach, which uses extracted features like the length of the branch, angles between branches and radius of the bronchus. We also intend to undertake a more robust evaluation on a larger number of patient datasets. This should aid in developing a normal lung model by efficiently modelling the airway tree topology and classification of cancerous regions.

4 Conclusions

We have described an automatic approach for the segmentation of lung areas from CT slices. The approach is robust and correlates well with manual segmentation results. However, the subsequent experiments are not conclusive in answering the thesis if patient orientation affects lung density estimation. We have shown limited correlation between prone and supine density estimation. Limitations of the current approach in combination with possible solutions have been discussed.

The segmented bronchus area gives a measure of the volume of air in the lungs and can be subtracted from total lung volume to gauge a better estimate of volume of lung tissue. Accurate segmentation of the bronchus is also vital for applications like virtual bronchoscopy. We have segmented the airway tree using a heuristic gain measure in a 3D region growing approach. Our results show a close correlation with manual segmentation and establish the use of gain as a reliable measure. A comparison of information gain based unsupervised image segmentation and boundary detection with alternate techniques can be found in [5].

References

1. S. Y. Hu, E. A. Hoffman & J. M. Reinhardt. “Automatic lung segmentation for accurate quantitation of volumetric x-ray ct images.” *IEEE Transactions on Medical Imaging* **20(6)**, pp. 490–498, 2001.
2. D. Aykac, E. Hoffman, G. McLennan et al. “Segmentation and analysis of the human airway tree from three-dimensional x-ray ct images.” *IEEE Transactions on Medical Imaging* **22**, pp. 940–950, 2003.
3. H. Singh, M. Crawford, J. Curtins et al. “Automated 3d segmentation of lung airway tree using gain-based region growing approach.” In *LNCS Medical Image Computing and Computer Assisted Intervention*. Springer-Verlag, 2004.
4. C. Revol-Muller, F. Peyrin, Y. Carrillon et al. “Automated 3D region growing algorithm based on assessment function.” *Pattern Recognition Letters* **23**, pp. 137, 2002.
5. H. Singh & R. Zwiggelaar. “Unsupervised image segmentation and boundary detection using information gain.” *The Irish Machine Vision and Image Processing Conference* 2004.

An image retrieval approach for thermal medical images

Gerald Schaefer[1], Shao Ying Zhu[2] and Bryan Jones[2]

[1]School of Computing and Technology, Nottingham Trent University, Nottingham, UK.
[2]Applied Computing, The Derbyshire Business School, University of Derby, Derby, UK.

Abstract

Past efforts on the automated processing on medical infrared images have typically focused on specialised applications like the detection of breast cancer. We propose the application of content-based image retrieval (CBIR) to thermal medical images. CBIR allows the retrieval of similar images based on features directly extracted from the image data. Hence, image retrieval for a thermal image that shows symptoms of a certain disease will provide visually similar cases which will usually also represent similarities in medical terms. The image features we propose for this purpose are a set of moment invariants of the grayscale thermal images.

Keywords: thermal medical imaging, content-based image retrieval, moment invariants.

1 Introduction

Image processing and pattern recognition techniques have been applied to infrared images for many years in astronomy and military applications for target recognition. Until recently, comparatively little effort has been expended in the automatic analysis of medical images. Several attempts have been made at automatic classification of thermal images, however all of them are specialised and hence restricted to a certain application e.g. the detection of breast cancer as in [1,2].

We propose the application of content-based image retrieval (CBIR) for medical infrared images as a generic approach for the automated processing of medical thermal images. CBIR allows the retrieval of relevant images based on a pre-defined similarity measure between image features. In terms of medical thermal imaging, images that are similar to a sample exhibiting symptoms of a certain disease or other disorder will be likely to show the same or similar manifestations of the disease. These known cases together with their medical reports will then provide a valuable asset for the diagnoses of the unknown case.

2 Thermal medical imaging

In recent years, there has been a resurgence of interest in the application of infrared thermal imaging in medicine due to improvements in camera technology and the promise of reduced costs [3]. Thermography captures the natural thermal radiation generated by an object at a temperature above absolute zero. The radiance from human skin is an exponential function of the surface temperature, which in turn is an indicator of the level of blood perfusion in the skin. Changes in blood perfusion may occur for a variety of reasons, such as inflammation, angiogenesis, and previous traumas. It is well known [4] that asymmetrical temperature distributions and hot and cold spots are strong indicators of an underlying dysfunction. Care is needed in interpreting thermograms since they are not specific and may reveal past traumas as well as current problems. It is non-invasive, radiation-free and complementary to anatomical investigations based on x-rays and three-dimensional scanning techniques such as CT and MRI and often reveals problems when the anatomy is otherwise normal. Computerised techniques of image processing and pattern recognition have been applied in acquiring and evaluating medical thermal images [5,6] which are important tools for clinical diagnostics.

3 Content-based image retrieval

Content-based image retrieval has been an active research area for more than a decade. The principal aim is to retrieve digital images based not on textual annotations but on features directly derived from the images data. These features are then stored alongside the image and server as an index. Retrieval is often performed in a query by example fashion where a query image is provided by the user. The image database is then searching through all images in order to find those with the most similar indices which are returned as the images most alike to the query image.

A large variety of features have been proposed in the CBIR literature [7]. However, in general they can be grouped into several categories: color features, texture features, shape features, sketch features, and spatial features. Often one or more feature types are combined in order to improve retrieval performance.

4 Content-based image retrieval for thermal medical images

We propose the use of CBIR for medical infrared images. One main advantage of using this concept is that it represents a generic approach to the automatic processing of such images. Rather than employing specialised techniques which will capture only one kind of disease or defect image retrieval when supported by a sufficiently large medical image database of both 'good' and 'bad' examples will provide those cases that are most similar to a given one. The query by example method is perfectly suited for this task with the thermal image of an 'unknown' case as the query image.

The features we propose to store as an index for each thermal image are invariant combinations of the moments of an image. Two-dimensional Cartesian moments m_{pq} of order $p+q$, of a density distribution function $f(x,y)$, are defined as

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy \qquad (1)$$

In terms of a digital image $g(x,y)$ of size $N \times M$ the calculation of m_{pq} becomes discretised and the integrals are hence being replaced by sums leading to

$$m_{pq} = \sum_{y=0}^{M-1}\sum_{x=0}^{N-1} x^p y^q g(x,y) \qquad (2)$$

Rather than mpq often centralised moments

$$\mu_{pq} = \sum_{y=0}^{M-1}\sum_{x=0}^{N-1} (x-\bar{x})^p (y-\bar{y})^q g(x,y) \quad (3)$$

with

$$\bar{x} = \frac{m_{10}}{m_{00}} \qquad \bar{y} = \frac{m_{01}}{m_{00}} \qquad (4)$$

It is well known that a small number of moments can characterise an image fairly well, it is equally known that moments can be used to reconstruct the original image [8]. Affine moment invariants which are independent of common operations such as scaling, translation, rotation and skew can be derived from centralised moments as [9]

$$\begin{aligned}
I_1 &= (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4 \\
I_2 &= (\mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{21}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2)/\mu_{00}^{10} \\
I_3 &= (\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2))/\mu_{00}^7 \qquad (5) \\
I_4 &= (\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 + 12\mu_{20}\mu_{11}^2\mu_{21}\mu_{03} + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - \\
&\quad 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - 8\mu_{11}^3\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + 9\mu_{20}\mu_{02}^2\mu_{21}^2 + 12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2)/\mu_{00}^{11}
\end{aligned}$$

Each thermal image is thus characterised by its four moment invariants. Image retrieval is performed by finding those images whose moments are closest to the ones calculated for a given query image. As a similarity metric (or rather distance measure) we use the Mahalanobis norm which takes into account different magnitudes of different components. The Mahalanobis distance between two moment invariant vectors V_1 and V_2 describing two thermal medical images I_1 and I_2 is defined as

$$d(I_1, I_2) = \sqrt{(V_1 - V_2)^T C^{-1} (V_1 - V_2)} \qquad (6)$$

where T denotes a transpose and C is the covariance matrix of the distribution of all Vs.

Ammer suggested, following a visiting fellowship at the University of Glamorgan [10], a number of standard views of patients. Clearly, the similarity would be defined as zero for comparisons between different views.

5 Experimental results

Figure 1 shows three thermal images of hands (in pseudocolour). The two pictures on the left are of healthy individuals while the hand on the right belongs to a sick person. We calculated the moment invariants for all three and the distances between the corresponding feature vectors. The distance between the two healthy individuals is 1.36 (despite the shift in position) while the distances between the healthy people and the sick person are 6.16 and 7.37 respectively thus confirming that the features introduced provide a promising avenue for exploring thermal medical images.

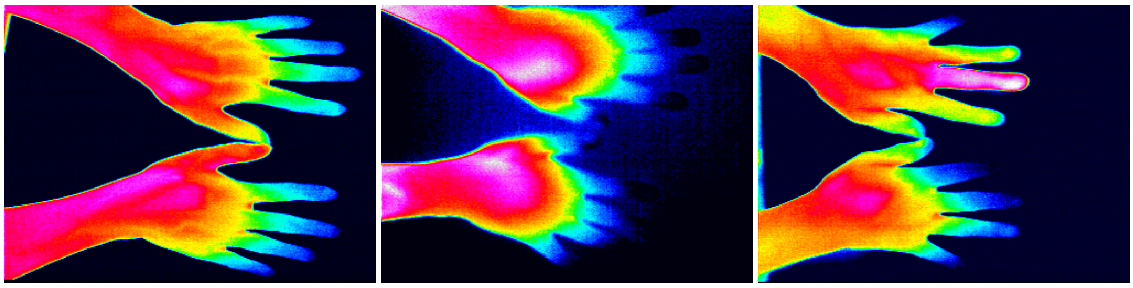


Figure 1. Thermal images of two healthy and one sick person.

6 Conclusions

We have proposed the application of content-based image retrieval to the domain of medical thermal images. Thermal images are characterised by a set of seven moment invariants; retrieval is performed by returning those images whose features are closest to the ones of a given query image.

Work is currently underway on building a database of medical thermal images comprising both images of 'normal' people [10] and cases of known symptoms of diseases which will enable to quantify retrieval results and evaluate future techniques.

Acknowledgements

The authors wish to thank the Nuffield Foundation for supporting this work.

References

1. H. Qi, "Breast cancer identification through shape analysis in thermal texture maps," *Proc. 2nd Joint EMBS-BMES Conf.*, vol. 2, pp. 1155-1156, Houston, 2002.
2. P.T. Kuruganti, H. Qi, "Asymmetry analysis in breast cancer detection using thermal infrared images," *Proc. 2nd Joint EMBS-BMES Conf.*, vol. 2, pp. 1129-1130, Houston, 2002.
3. B.F. Jones, "A re-appraisal of the use of infrared thermal image analysis in medicine", *IEEE Trans. Medical Imaging*, Vol. 17, no. 6, pp. 1019-1027, 1998.
4. S. Uematsu, "Symmetry of skin temperature comparing one side of the body to the other", *Thermology*, vol. 1, pp. 4-7.
5. P. Plassmann and E.F.J. Ring, "An open system for the acquisition and evaluation of medical thermological images", *Eur. J. Thermology*, vol. 7, no. 4. pp. 216-220,1997.
6. B. Wiececk, S. Zwolenik, a. Jung, J. Zuber, "Advanced thermal and radiological image processing for clinical diagnostics", *Proc. 1st Joint IEEE BMES-EMBS Conference*, Atlanta, 1999.
7. A. Smeulder, M. Worring, S. Santini, A. Gupta, R. Jain, "Content-Based Image Retrieval at the End of the Early Years", *IEEE Trans. PAMI*, vol. 22, no. 12, pp. 1349-1380, 2000.
8. M. Hu, "Visual pattern recognition by moment invariants", *IRE Trans. Inform. Theory*, vol. 8, pp. 179-187, 1962.
9. J. Flusser and T. Suk, "Pattern recognition by affine moment invariants", *Pattern Recognition*, Vol. 26, No. 1, pp.167-174, 1993.
10. EPSRC Grant BG/R50134/01.

MAP model of simultaneous segmentation and registration

Chen Xiaohua[a], Michael Brady[a] and Daniel Rueckert[b]

[a]Robotics Research Group, University of Oxford, Oxford, UK
[b]Department of Computing, Imperial College London, London, UK

1 Introduction

Although segmentation and registration are usually considered separately in medical image analysis, they can benefit a great deal from each other. Until recently, little research has been done to simultaneously estimate the segmentation and registration problems in a single framework to make the two problems' solutions facilitate each other. Yezzi [1] and Wyatt [2] used a number of methods for interleaving segmentation and registration, but both methods have limitations. In this paper, we present a novel maximum a posteriori (MAP) model for simultaneous segmentation and registration (SSR) and applied it on brain MR images.

2 Method

We have two images I and J, and we assume J corresponds to some unknown geometric transformation of reference image I. "Segmentation", or the labelling of each pixel to one tissue type can be regarded as a model of the underlying anatomy. I and J can be interpreted as a realization of a random process that corrupts the "Segmentation", e.g. by Gaussian noise. The problem can be formulated as follows: given image I and J, we wish to simultaneously estimate the label fields f of the images and recover the geometric transformation T that registers the two images. The MAP estimation is to find f and T to maximize $P(f, T|I, J)$.

2.1 Hidden vector field

We assume that there are M regions; discrete label $f(\vec{r})$ indicates to which region pixel $\vec{r} = (x_r, y_r)$ belongs. As noted above, the relationship between an image and its segmentation can be defined in terms of a Gaussian noise distribution: $P(I|f) = \prod_{\vec{r}\in\Omega_I} v_{f(\vec{r})}(\vec{r})$. where Ω_I is the lattice of sites of I and each M-vector $\vec{v}(\vec{r})$ is defined by: $v_k(\vec{r}) = \sqrt{\frac{\gamma}{\pi}}\exp[-\gamma|I(\vec{r}) - \theta_k|^2]$. Here, γ and $\vec{\theta}$ are image parameters that depend on the noise variance and mean intensity value of each class. One can find the optional estimate for f by applying a classical MRF model and Bayesian MAP estimation [3].

Marroquin [4] proposed a different probabilistic model for the generation of label field to overcome the difficulties with classical MRF models like sensitive to noise and initialization. Instead of the conventional 1-step procedure, he proposed a 2-step probabilistic model, with an additional hidden Markov random vector field p: each vector $\vec{p}(\vec{r})$ indicates the probability the pixel $\vec{r}$ belongs to one of the regions given the intensity of that pixel, and it takes values on the *M*-vertex simplex S_M: $S_M = \{\vec{u} \in \mathcal{R}^M : \sum_{k=1}^{M} u_k = 1, u_k \geq 0, k = 1, ..., M\}$. Then, the optimal estimate for f can be calculated from p.

2.2 General framework

We incorporate the Markov random vector field p into our framework. To obtain the optimal estimator f^* for the label field and T^* for transformation, we follow the steps:

1. Find the MAP estimators p^*, T^* for p, T: $p^*, T^* = \arg\max_{p\in S_M^N, T} P(p, T|I, J)$
2. Determine $f^*(\vec{r}) = \arg\max_{f(\vec{r})} P(f|p = p^*, I) = \arg\max_k p_k^*(r)$

The first step itself is a 2-step procedure, in which the best T is found given the current estimate for p, then the best estimate for p is found, given the current estimate for transformation T:

1. Find an initial estimate $\bar{p}$ for p by individual segmentation;

2. Repeat until convergence or often enough:

   (a): Set $\bar{T} = \arg\max_T P(T|\bar{p}, I, J)$

   (b): Set $\bar{p} = \arg\max_p P(p|\bar{T}, I, J)$

We now analyze step (a). We consider image I to be the reference image, transformation T to be a spatial mapping from I to J. Using Bayes' rule, we have: $P(T|\bar{p}, I, J) \propto P(I|\bar{p})P(J|\bar{p}, T)P(T)$.

In order to maintain spatial coherence and smoothness, the transformation $T(\vec{r})$ may be required to be similar to its value at the spatial neighbors. We assume a Gibbs distribution on the expected deformations: $P(T) = \exp(-E(T))$, where $E(T)$ is T in the form of an energy. The likelihood of the observations can be rewritten: $P(J|\bar{p}, T) = \prod_{\vec{r}\in\Omega_I} P(J(T(\vec{r}))|\bar{p})$. For a Gaussian noise distribution, we have:

$$p(J(T(\vec{r}))|\bar{p}) = \sum_{k=1}^{M} w_k(T(\vec{r}))\bar{p}_k(\vec{r}) = \vec{w}(T(\vec{r})) \cdot \bar{\vec{p}}(\vec{r}) \ .$$

where $w_k(T(\vec{r})) = \sqrt{\frac{\gamma}{\pi}}\exp[-\gamma|J(T(\vec{r})) - \theta_k)|^2]$. Finally we get: $P(T|\bar{p}, I, J) \propto \exp[-U(T)]$. where $U(T) = -\sum_{\vec{r}\in\Omega_I} \log(\vec{w}(T(\vec{r})) \cdot \bar{\vec{p}}(\vec{r})) + E(T)$. Step (a) is equivalent to minimizing of $U(T)$.

For step (b): $P(p|\bar{T}, I, J) \propto P(I|p)P(J|p, \bar{T})P(p)$. Since p is Markovian, $P(p)$ can be expressed as $P(p) = \exp(-V(p))$, where, for example, $V_{\vec{r}\vec{s}}(\vec{p}(\vec{r}), \vec{p}(\vec{s})) = \lambda|\vec{p}(\vec{r}) - \vec{p}(\vec{s})|^2 = \lambda\sum_{k=1}^{M}(p_k(\vec{r}) - p_k(\vec{s}))^2$. where λ is a positive parameter, and $< \vec{r}, \vec{s} >$ are neighboring sites in Ω_I.

So, we get: $P(p|\bar{T}, I, J) \propto \exp[-U(p)]$. with $U(p) = -\sum_{\vec{r}\in\Omega_I} \log(\vec{v}(\vec{r}) \cdot \vec{p}(\vec{r})) - \sum_{\vec{r}\in\Omega_I} \log(\vec{w}(\bar{T}(\vec{r})) \cdot \vec{p}(\vec{r})) + \sum_C V_C(p)$.

Step (b) is then equivalent to minimizing of energy $U(p)$. we use iterative gradient descent method for optimization of both T and p. Here, $\vec{p}(\vec{r})$ must be projected back into S_M.

2.3 The representation of transformation

For the case where T is a rigid transformation, we represent it by a rotation matrix A and a translation vector $\vec{c}$: $T(\vec{r}) = A\vec{r} + \vec{c}$. In two dimensions, the rotation matrix A depends upon a single angle α. Since for a rigid registration, all the pixels undergo the same transformation, it is sufficiently smooth for $E(T)$ to be dropped, so we can set it to zero. By minimizing $U(T)$, we get parameters $\alpha, \vec{c}$ to represent $\bar{T}$.

For non-rigid transformation, we represent T using a combination of a global transformation and a local transformation: $T = T_{global} + T_{local}$. The global transformation is represented by a rigid transformation, while for and the local transformation, we use a B-spline based FFD model [5].

For T_{global}, the $E(T)$ can be dropped, and for T_{local}, we may use a 5 pixels neighborhood clique C: $E(T) = \sum_C V_C(T)$ where

$$\begin{aligned}\sum_C V_C(T(\vec{r})) &= 2(T(x_r, y_r + 1) + T(x_r - 1, y_r) - T(x_r, y_r) - T(x_r - 1, y_r + 1))^2 \\ &+ (T(x_r - 1, y_r) + T(x_r + 1, y_r) - 2T(x_r, y_r))^2 \\ &+ (T(x_r, y_r - 1) + T(x_r, y_r + 1) - 2T(x_r, y_r))^2 \ .\end{aligned}$$

3 Results

In this section, we use brain MR images to illustrate the performance of our approach of simultaneous segmentation and registration presented above.

In order to compare the algorithms, we need to establish a performance index, which should be objective and quantitative. We propose the following performance index ξ: $\xi_k = \frac{2V_{GPk}}{V_{Pk}+V_{Gk}}$ where V_{GPk} denotes the total

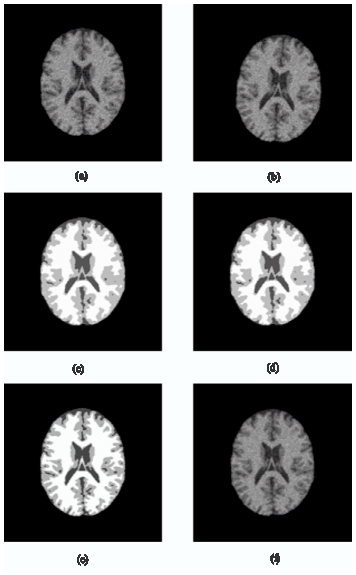


Figure 1. Simultaneous segmentation and registration (SSR) experiment on rigid registration. (a)reference image I. (b)floating image J. (c)segmentation result of image I without SSR. (d)segmentation result with SSR. (e)segmentation ground truth in image I domain. (f)transformation of floating image to reference image using rigid registration.

tissue class	CSF	GM	WM
single	0.893	0.869	0.930
SSR	0.894	0.874	0.937

Table 1. performance index comparison between single and SSR segmentation for experiment1

number of pixels that were correctly assigned to class k by a given procedure; V_{Pk} is the total (correct + incorrect) number of pixels belonging to class k by this procedure and V_{Gk} denotes the total number of pixels belonging to class k in the ground truth. Higher performance index indicates better segmentation result here.

Our initial experiments are on brain MR images provided by Brainweb [6]. We segment the images into 3 tissue classes in the brain: cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM). Our first experiment is on images with known rigid transformation between them. We make a transformation on zero noise brain image with 5.00 degree rotation and 11.0 pixels translations in both directions. Various amount of zero-mean Gaussian white noise are independently added to the original and transformed images to produce the observed images I and J. The test images, together with their single and SSR results are shown in Fig. 1. We compare these two results with our performance index: as can be seen from Table 1, in each tissue class, the performance index of SSR is always higher than that of the single segmentation. The recovered transformation is 5.03 degree and the translations in two direction are 10.7 and 10.4 pixels.

Our second experiment is on images with an unknown non-rigid transformation: we take two different slices of brain MRI as I and J; the image experimental results are shown in Fig. 2. We can see from the performance index in Table 2 that the SSR gives more correct classification for pixels in each class than the single segmentation. The recovered transformation is represented by transforming the floating image into the reference image domain, since no real transformation ground truth can be provided here, we can only get a visually qualitative impression of our registration result. We aim to study how to evaluate non-rigid registration methods later.

tissue class	CSF	GM	WM
single	0.882	0.888	0.939
SSR	0.890	0.892	0.942

Table 2. performance index comparison between single and SSR segmentation for experiment2

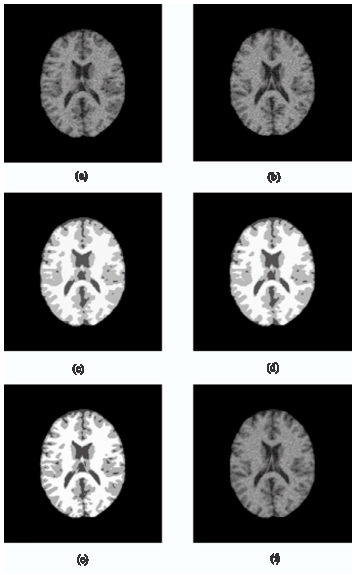


Figure 2. Simultaneous segmentation and registration (SSR) on non-rigid registration. (a)reference image I. (b)floating image J. (c)segmentation result of image I without SSR.(d)segmentation result with SSR. (e)segmentation ground truth in image I domain. (f)transformation of floating image to reference image using non-rigid registration.

4 Discussion

In this paper, we have developed a framework to achieve simultaneously segmentation and registration to make the two problems' solutions facilitate each other. We use a hidden Markov measure vector field to make the interactions between these two problems possible. It is used for segmentation step to label each pixel with highest probability of certain tissue type and for registration step to act as a key element in similarity measure. For reason of space, we have only illustrated this framework to brain MR images for both rigid and non-rigid registration cases, with promising results for both segmentation and registration: the segmentation results achieved by fusion of the images performs better than the segmentation results got from single image, meanwhile, using the segmentation results also produces a good registration.

References

1. A. Yezzi & L. Zollei. "A variational framework for integrating segmentation and registration through active contours." *Medical Image Analysis* **7**, pp. 171–185, 2003.
2. P. P. Wyatt & J. A. Noble. "MAP MRF joint segmentation and registration of medical images." *Medical Image Analysis* **7**, pp. 539–552, 2003.
3. Y. Zhang, M. Brady & S. Smith. "Segmentation of brain MR images through a hidden Markov random fi eld model and the expectation maximization algorithm." *IEEE Trans. Medical Imaging* **20**, pp. 45–57, 2001.
4. J. L. Marroquin & S. B. E A Santana. "Hidden Markov measure fi eld models for image segmentation." *IEEE Trans. Pattern Analysis and Machine Intelligence* **25**, pp. 1380–1387, 2003.
5. D. Rueckert, L. I. Sonoda, C. Hayes et al. "Nonrigid registration using free-form deformations: Application to breast MR images." *IEEE Trans. Medical Imaging* **18**, pp. 712–721, 1999.
6. C. A. Cocosco, V. Kollokian, R. Kwan et al. "BrainWeb: Onlie interface to a 3D MRI simulated brain database." *Neuroimage* **5**, pp. 425, 1997.

Registration of Tomographic Images of the Optic Disc

Jim Ivins and Andrew Turpin[1]

Department of Computing, Curtin University of Technology, Western Australia

Abstract. Glaucoma is a degenerative disease of the eye, and the largest cause of blindness in the industrialised world. It is typically associated with raised Intra-Ocular Pressure (IOP). The Heidelberg Retina Tomograph (HRT) is an imaging device capable of accurately measuring the depth of the retinal surface in the posterior half of the eye. HRT images of glaucomatous eyes typically show degeneration of the retina in an area called the optic disk, where the optic nerve leaves the eye. The degeneration tends to get worse as IOP increases and the disease progresses. This paper describes a technique for registering multiple HRT images of the same subject at different levels of IOP, taken during experiments to simulate glaucoma. The resulting fully registered images are currently being used to study the underlying aetiology of glaucoma.

1 Introduction

Glaucoma is a degenerative disease of the eye. Symptoms typically begin with some loss of peripheral vision – the patient can no longer see 'out of the corner of the eye'. Over a period of years the size of this visual defect gradually increases, often resulting in total blindness. Glaucoma currently affects roughly 150 000 people in Australia (0.75% of the population), but this figure is predicted to exceed 300 000 by 2030, with similar rates in other industrialised countries [1]. Although the clinical manifestation is well known, the underlying causes of glaucoma are less well understood. Consequently, there is an urgent need to understand the aetiology of the condition so that better prevention, detection and treatment methods can be devised.

Figure 1 shows the retinal surface at the optic disc or 'blind spot'. In this area blood vessels and nerves from the retina come together and pass through a mesh of connective tissue called the *lamina cribrosa* to form the optic nerve. The result is an obvious depression of the internal surface of the retina – a feature called the *optic disc*. One of the main clinical signs of glaucoma is an abnormally high Intra-Ocular Pressure (IOP) exerted by the fluid (vitreous humor) within the eye. Several researchers have suggested that high IOP somehow affects the nerve fibres and/or blood vessels as they pass through the lamina cribrosa (see [2] for a review). This leads to gradual nerve death and hence loss of vision. As the nerves die, the optic disc becomes greatly enlarged. Unfortunately, the exact details of the process are poorly understood.

The goal of this work is to study the effects of changing IOP on the retina, using image-processing techniques to examine the retinal surface at the optic disc [3]. As a first step towards this goal, this paper describes the method used to register images of the optic disc taken at various levels of artificially raised or lowered IOP. Section 2 describes the image acquisition method. Section 3 describes the registration method. Section 4 describes some experimental work using these methods. Finally, Section 5 draws some conclusions.

2 Retinal Tomography

The Heidelberg Retina Tomograph (HRT) is a confocal laser-scanning microscope designed to acquire depth images of the internal surface of the eye via the pupil [4]. The result is a two-dimensional matrix of depth measurements that are reproducible to within 20 microns. This topography image consists of 256×256 individual depth measurements that are absolutely scaled for the individual eye. Figure 1 (LHS) shows a typical topography image of the optic disc of a normal rabbit eye. Pixel intensity is a measure of surface depth, such that white pixels are deeper than black pixels. The corresponding surface plot is also shown (RHS).

The main clinical use of the HRT is in routine monitoring of the glaucomatous eye. However, researchers are using the HRT to describe topographic changes in the retina caused by increasing or decreasing the IOP under laboratory conditions [5]. The first step in analysing the data from this work is image registration. The HRT is shipped with software intended for use by clinicians to analyse human retinal images. However, in laboratory work with non-human subjects this software was often unable to register images. Initially this was though to be partly due to the large changes seen in the images, and partly because the optic disc of the rabbit has a different structure from its human equivalent. It was therefore necessary to develop an alternative registration method to support the laboratory work. The fundamental assumption when registering clinical HRT images is that features in the peripheral part of the image do not change. In other words, while the optic disc may be deforming under pressure, the surface in the peripapillary region is unchanged. However, as the experimental work described in Section 4 revealed, this is not the case.

[1] Email: {jim, andrew}@cs.curtin.edu.au.

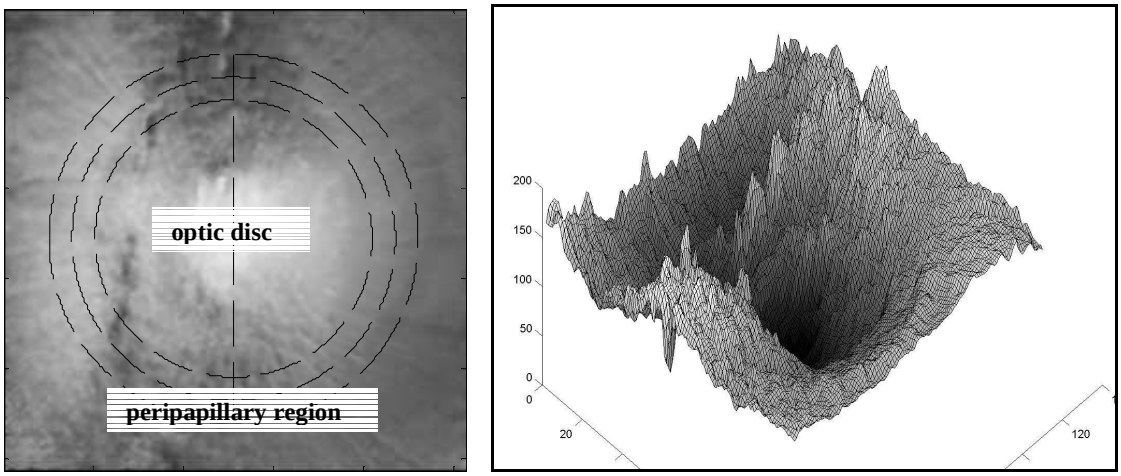


Figure 1. (LHS) HRT data of a normal rabbit eye shown as an image. Note the reference rings, which are the basis of the registration method. **(RHS)** The same data shown as a surface map.

3 Registration Method

Prior to registration each image is filtered with a difference of Gaussians. A small Gaussian mask eliminates high-frequency changes such as noise. *A large mask eliminates low-frequency changes such as depth gradients across the image due to tilting of the eye relative to the HRT.* A difference of large and small Gaussians acts as a band-pass filter, producing an image in which noise is suppressed while edges are highlighted. This helps to prevent the registration algorithm from becoming trapped in local minima. Given an image containing 256-by-256 pixels, a suitable pair of Gaussian kernels span 7 and 13 pixels respectively, and cover ±2 standard deviations.

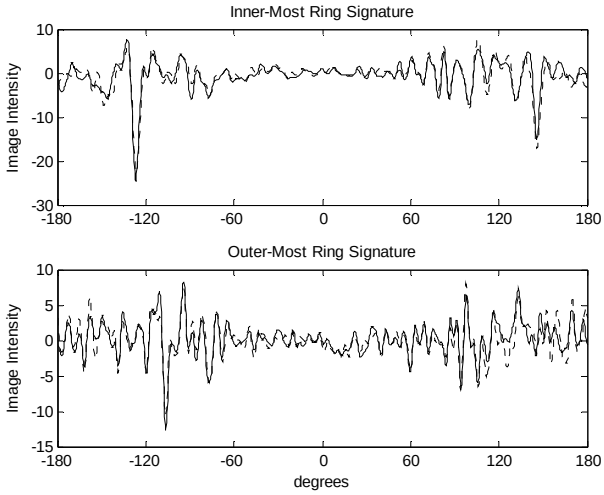


Figure 2. Retinal signatures around reference rings.

After pre-processing, a set of circular reference rings is selected from the peripapillary region of the reference image by the user specifying the centre of the optic disc, and inner and outer ring radii (see Figure 1 LHS). Each reference ring represents a profile or 'signature' of the retinal surface extending 360º around the optic disc as shown in Figure 2 (solid lines). The corresponding signatures from the unregistered image (dashed lines) are aligned with the reference image by transforming the rings using a method suggested by Ivins et al [6]. The state of the unregistered image can then be described by just a few parameters, such as translation and rotation, which control the configuration of the rings in the image. The parameters of the transformation are varied to minimise differences between the retinal signatures from the reference image, and the equivalent signatures from the unregistered image. This is achieved using a non-linear least-squares minimization algorithm. The expected motion determines the choice of parameters used to construct the model.

The most important change between images is that due to experimental manipulation of the IOP. Obviously, this change must not be affected by registration. In addition, the following sources of change occur in the experiments. Note that movements of the HRT device are equivalent to movements of the eye in terms of their effects on the retinal images.

1. Translation of the eye in three dimensions. This results in horizontal and vertical displacements of the image, which can be corrected using image translation. Translation in the image plane does not affect depth values. However, translation of the eye orthogonal to the image plane affects depth values throughout the image. The HRT produces images that are absolutely scaled, but it is still necessary to adjust the mean depth of each unregistered image. This type of translation will also produce scaling in the image plane.

2. Rotation of the eye in three dimensions. The result is a combination of rotation in the image plane, which does not affect depth values, and shearing of the image with corresponding changes in depth. Depth values increase in some parts of the image and decrease in others.
3. Changes to the focal settings of the HRT. These can be ignored if the machine parameter is not altered during an experiment.

3.1 Non-Linear Least-Squares Minimization

If the registration method is to avoid over fitting the changes with too many parameters then the components must be approximated by just a few degrees of freedom. An affine transformation is sufficient to model the positional changes seen in the images. The transformation includes translation, rotation, scaling and two kinds of shearing (deformation). However, it does not affect the image depth values in any way, so a separate correction must be applied afterwards.[2] A point (x, y) in the original reference ring is warped to produce an equivalent point in an unregistered image by translation (t_x, t_y), and an affine matrix **A**:

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{bmatrix} 1+a_{xx} & a_{xy} \\ a_{yx} & 1+a_{yy} \end{bmatrix} \begin{pmatrix} x-c_x \\ y-c_y \end{pmatrix} + \begin{pmatrix} t_x+c_x \\ t_y+c_y \end{pmatrix} \text{ where } \begin{pmatrix} c_x \\ c_y \end{pmatrix} = \frac{1}{M}\sum \begin{pmatrix} x \\ y \end{pmatrix}$$

During this transformation, the centroid (c_x, c_y) of the model (computed from the co-ordinates of the M samples in the profiles) must coincide with the origin of the co-ordinate system. It is therefore necessary to subtract the centroid before transforming the rings, and add it on afterwards.

The combined parameters $\mathbf{p} = (p_1, p_2, \ldots, p_N)$ of the model (where $N = 6$) are iteratively adjusted to find the optimal registration between the rings in the unregistered image and those from the reference image. This is achieved using a non-linear least-squares minimization algorithm. A discrete error function $E(\mathbf{p})$ is minimized which is a sum of squares:

$$E(\mathbf{p}) = \frac{1}{2}\mathbf{e}(\mathbf{p})^T\mathbf{e}(\mathbf{p}) = \frac{1}{2}\sum_{i=1}^{M} e_i(\mathbf{p})^2 \text{ where } e_i(\mathbf{p}) = I_1(\mathbf{s}_1) - I_2(\mathbf{s}_2) = I_1(\mathbf{s}_1) - I_2(\phi(\mathbf{s}_1, \mathbf{p}))$$

The co-ordinates $\mathbf{s}_1$ from the rings in the unregistered image I_1 are warped by some function $\phi(\mathbf{p})$ to produce equivalent co-ordinates $\mathbf{s}_2$ in the reference image I_2. The M measurements e_i are the intensity differences at corresponding co-ordinates along the two reference rings.

Finally, after applying the affine transformation the depth of the newly registered image is brought into alignment with the reference image using a linear least-squares method to compute a constant offset and two first-order gradient parameters. (In the experimental work these were found to be very small).

Despite the complexity implied by using a non-linear method, certain characteristics of the least-squares problem can be exploited to improve efficiency. In particular, the gradient and Hessian matrices (which must be calculated repeatedly) have special structures. If the M-by-N Jacobian matrix of the residual $\mathbf{e}(\mathbf{p})$ is denoted as $\mathbf{J}(\mathbf{p})$ then the gradient $\nabla\mathbf{e}(\mathbf{p})$ and the Hessian $\mathbf{H}(\mathbf{p})$ of $\mathbf{e}(\mathbf{p})$ are defined as:

$$\nabla\mathbf{e}(\mathbf{p}) = \mathbf{J}(\mathbf{p})^T\mathbf{e}(\mathbf{p}) \qquad \mathbf{H}(\mathbf{p}) = \mathbf{J}(\mathbf{p})^T\mathbf{J}(\mathbf{p}) + \mathbf{Q}(\mathbf{p})$$

The Hessian matrix has the property that when the residual tends to zero, then as $\mathbf{p}$ approaches the solution, the second-order term $\mathbf{Q}(\mathbf{p})$ also tends to zero. This fact can be exploited to minimize the least-squares error using the Levenberg-Marquardt method, an implementation of which is included in the MatLab Optimisation Toolbox [7]. The minimization problem is solved most efficiently if a function is provided to compute the partial derivatives of the residual (the Jacobians). Given co-ordinates $\mathbf{s}_1$ from reference rings in an image I_1 and co-ordinates $\mathbf{s}_2 = \phi(\mathbf{s}_1, \mathbf{p})$ from the rings in an unregistered image I_2 the error between the two is defined as:

$$E = \frac{1}{2}\sum_{i=1}^{M} e_i(\mathbf{p})^2 \text{ where } e_i(\mathbf{p}) = I_1(\mathbf{s}_1) - I_2(\mathbf{s}_2)$$

The partial derivatives are as follows:

$$\frac{\partial E}{\partial p} = \sum_{i=1}^{M} e_i(\mathbf{p})\frac{\partial e_i}{\partial \mathbf{p}} = \sum_{i=1}^{M} e_i(\mathbf{p})\frac{\partial e_i}{\partial \mathbf{s}_2}\cdot\frac{\partial \mathbf{s}_2}{\partial \mathbf{p}} = \sum_{i=1}^{M} e_i(\mathbf{p})\nabla I_2(\mathbf{s}_2)\cdot\frac{\partial \phi}{\partial \mathbf{p}}$$

Values for the residual $e_i(\mathbf{p})$ are easily computed by subtracting the appropriate image intensities. The image gradient ∇I_2 at co-ordinates $\mathbf{s}_2$ is found using finite differences and bi-cubic interpolation. The final term in the summation gives the position dependence of the rings on the transformation parameters. Appropriate formulae for these partial derivatives must be obtained explicitly. There are six partial derivatives for each co-ordinate (one for each parameter), but fortunately the values of these derivatives are easy to compute.

[2] Originally the intention was to use a 3-D model with three translation and three rotation parameters. However, band-pass filtering obliterates the depth information necessary for this approach. Unfortunately the band-pass filtering is essential for the minimisation algorithm, which would otherwise be distracted by local minima.

4 Experimental Work

The registration method was implemented using MatLab. To register a 256-by-256 pixel image takes 3–5 seconds on an unremarkable PC (a Celeron 450 processor with 128MB of RAM running Windows 2000). The method is currently being used to study the deformation of the laminar cribrosa in artificially induced glaucoma. **Aim.** The aim of the experiment was to study the biomechanics of the optic disk in glaucoma. **Method.** Six rabbits had IOP artificially increased to provide a sequence of known pressure gradients across the lamina cribrosa. At each pressure gradient, an image centred on the disk with a 15° field of view was obtained using the HRT. The registration algorithm described in this paper was used to register the images and hence make movies of the retinas as IOP increased. **Results.** Figure 3 shows some typical registered images obtained from one subject at different IOP levels. The corresponding movies reveal that there are major changes in the retinal surface due to pressure, not only in the optic disk, but also in the peripapillary region [8]. This makes registering the images difficult, because there is no point on the retinal surface that is consistently unaffected by altering the pressure gradient. This may explain why off-the-shelf software failed to register the images. **Conclusion.** The assumption, that there is an area of the retina that does not change when the pressure gradient over the lamina changes is false in rabbit eyes. This implies that image processing techniques that assume an unchanging peripapillary region must be used with caution when attempting to measure changes in images obtained with HRT. It is an open question whether this finding is also true in human eyes.

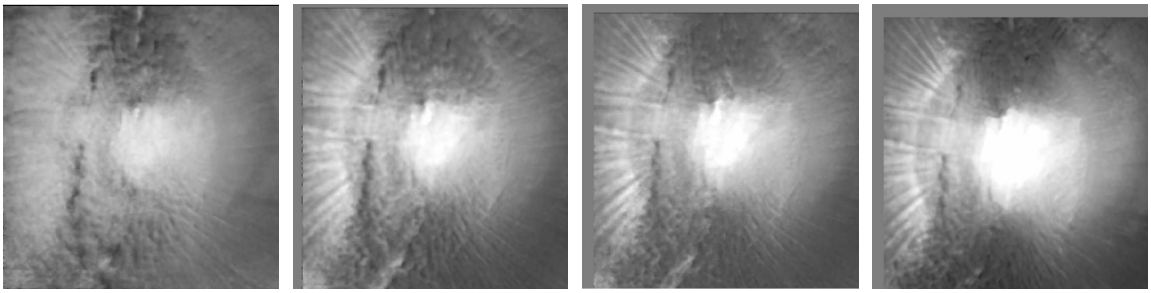


Figure 3. A registered image sequence of the optic disc (light pixels are deeper than dark ones). Note the effects of increasing pressure (from left to right). The gray borders are due to the registration method.

5 Conclusions

The registration method described in this paper is based on a low parameter model driven by non-linear least squares minimization, giving it two main advantages over existing HRT registration software. First, the method can be used to register images even when there are obvious changes in the peripapillary region. Second, different mathematical models of image registration can be 'plugged in' to the software with ease – for example, to replace the affine transformation with a more (or less) sophisticated one.

Acknowledgements

Dr W Morgan at the Lion's Eye Institute in Perth, Western Australia, supplied the images used in this paper.

References

[1] E. Rochtchina, P. Mitchell. "Projected number of Australians with glaucoma in 2000 and 2030." *Clinical and Experimental Ophthalmology,* **28**, pp 146–148, 2000.
[2] G. A. Cioffi. "Three common assumptions about ocular blood flow and glaucoma." *Survey of Ophthalmology,* **45** (supplement 3), S325–S331, 2001.
[3] A. Turpin, J. Ivins, W. Morgan. "Dynamics of the Optic Disk In a Known Pressure Gradient Measured With Scanning Laser Ophthalmoscopy." Australasian Ophthalmic and Visual Sciences Meeting (Melbourne, Australia), 2003.
[4] Heidelberg Engineering GmbH, 69221 Dossenheim, Germany http://www.heidelbergengineering.com
[5] W.H. Morgan, B.C. Chauhan, D-Y Yu, et al. "Optic Disc Movement with Variations in Intra-ocular and Cerebrospinal Fluid Pressure." *Investigative Ophthalmology & Visual Science,* **43**, pp 3236–3242, 2002.
[6] J. Ivins, J. Porrill, J. Frisby. "A deformable model of the human iris driven by non-linear least-squares minimisation." *Sixth International Conference on Image Processing and its Applications (Dublin, Ireland):* **1**, pp 234–238, 1997.
[7] The Mathworks, Inc: 2000. "Optimisation Toolbox User's Guide version 2 (for use with MatLab)." The MathWorks Inc.
[8] Sample movies of one subject are available from http://biomed.cs.curtin.edu.au/projects.html

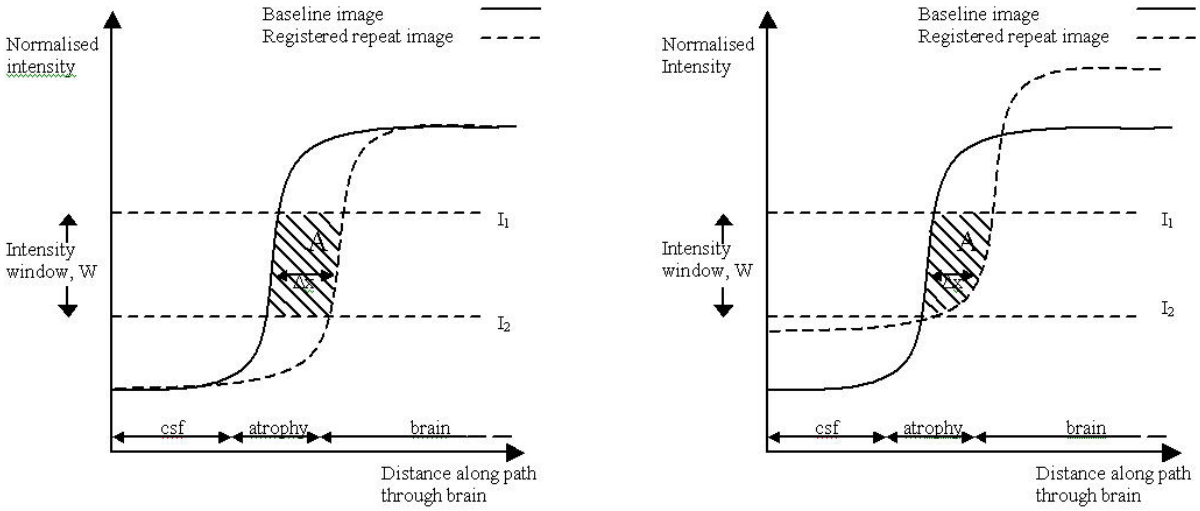
Correcting differential intensity inhomogeneity

Emma B Lewis[a]*and Nicholas C Fox[a]

[a]Dementia Research Centre, Institute of Neurology, 8-11 Queen Square, London WC1N 3BG, UK

1 Introduction

Serial imaging is increasingly important in dementia: for diagnosis, for monitoring disease progression and for measuring drug modification effects in clinical trials and treatment management. Longitudinal measures, e.g. atrophy quantification, are especially sensitive to scan quality. One common scan artifact is intensity inhomogeneity (bias): a slowly changing, smooth spatial variation in signal intensity within the scan. This artifact may, for example, result in tissue in one part of the brain having a systematically lower signal intensity than similar tissue in another part of the brain. This effect can be in the order of 20%. There are several causes of intensity inhomogeneity: inhomogeneity of the magnetic field, B_0, of the MR system; inhomogeneity of the radiofrequency (RF) pulse generated by the oscillating secondary magnetic field, B_1, or non-uniform sensitivity of the receiver coils used to detect the MR signal. Intensity inhomogeneity creates several problems for subsequent image processing. Segmentation of an image into its constituent tissue types becomes more difficult - especially for automated, intensity thresholding based techniques. In particular, differences between the bias fields of the longitudinal image pair confound longitudinal image processing techniques: registrations, especially non-linear, may be affected, and intensity based quantification techniques will be impaired. Figure 1 illustrates the impact of differential bias on a well known intensity based atrophy measure, the Brain Boundary Shift Integral (BBSI) [1].



(a) (b)

Figure 1. Idealised intensity profiles in one dimension of a longitudinal scan pair across a csf - brain boundary for (a) the case where there is negligible differential intensity inhomongeneity and (b) the case where there is significant differential bias.

Figure 1(a) illustrates the idealised intensity profile in 1d across a cerebrospinal fluid to brain boundary where Δx represents the distance moved by an anatomical point, the *boundary shift* and is given in Eq. 1 thus:

$$\Delta x \approx A/W \tag{1}$$

where A is the clipped boundary volume and W is the selected intensity window which should be contained within the intensity transition of the boundary.

The BBSI approximates the volume of atrophy in 3D as the total boundary shift as in Eq. 2.

*Email: elewis@dementia.ion.ucl.ac.uk

$$BBSI = \sum_{boundary} [K/(I_2 - I_1)[clip(u_b(x), I_2, I_1) - clip(u_r(x), I_2, I_1)] \quad (2)$$

where $clip(a(x), I_1, I_2) = I_2$ if $a(x) > I_2$, $clip(a(x), I_1, I_2) = a(x)$ if $I_1 < a(x) < I_2$ and $clip(a(x), I_1, I_2) = I_1$ if $a(x) < I_1$ and K = voxel volume.

Figure 1(b) illustrates the same profile for the case where there is a significant differential bias between the longitudinal scans. It can be seen that the BBSI calculated from Eq. 2 will not give an accurate representation of atrophy.

Current bias correction techniques are either modified acquisition protocols or post-processing techniques and correct absolute bias on single scans. Modified acquisition protocols include scanning the object twice using two different coils [2], and measuring the RF field during the scan [3]. Post-processing techniques include modelling intensity variation in regions of assumed homogeneity provided by tissue type segmentations [4]. More sophisticated techniques combine segmentation and intensity correction in a single algorithm e.g. [5]. The most commonly used bias correction technique, N3 [6], obtains the spline based bias field by iteratively sharpening the joint intensity histogram. However, modified acquisition protocols result in lengthened scan acquisition times, and cannot be used for retrospective correction, while current post-processing techniques assume models for tissue-type or bias field, or assume frequency domain separability of anatomy and bias.

The technique described here corrects *differential* bias on an image pair to improve measurement of pathological change from longitudinal scans, thereby avoiding the need for the assumptions required when correcting absolute bias. It identifies the differential bias field by applying simple filters to the longitudinal pair's difference image, which consists of differential bias, anatomical change (e.g. atrophy), registration error and noise, to obtain the relatively large scale differential bias field which is then used to correct the longitudinal scan pair for subsequent image processing.

2 Methods

2.1 Theory

Baseline and repeat brain regions are initially segmented using a semi-automated technique. The repeat image is then rigidly registered, with rescaling, to the baseline image and resliced, and the repeat brain region resliced using the same transformation. Intensities are normalised to the mean intensity of the interior brain region. Image formation is modelled with a multiplicative bias field and additive white Gaussian noise. Calculating the difference image of the log-transformed baseline and repeat images gives Eq. 3:

$$log(v_b(x)) - log(v_r(x)) = log(b_b(x)/b_r(x)) + log(u_b(x)) - log(u_r(x)) + n'(x) \quad (3)$$

where $v_b(x)$ and $v_r(x)$ are the measured signal in the baseline and repeat images respectively, $u_b(x)$ and $u_r(x)$ are the true signal in the baseline and repeat images, $b_b(x)$ and $b_r(x)$ are the bias fields in the baseline and repeat images and $n'(x)$ is approximately white Gaussian noise (modified by signal and bias). The term $log(u_b(x) - log(u_r(x))$, consists of anatomical change, i.e. atrophy, and registration error. Since the images are intra-subject, longitudinal scans then registration error should be small, and assuming that there have been no regional changes to the intensity of grey or white matter, this term will have relatively small scale structure. The noise term will similarly be of small scale structure. The term $log(b_b(x)/b_r(x))$ represents the differential bias field; this is of relatively large-scale structure [6]. A median filter is then applied to Eq. 3. The median filter removes Gaussian noise [7] and erases structure of size less than half that of the kernel [8] which should result in atrophy and registration error being removed by the filter, to leave the differential bias field. A kernel of 11x11x11 voxels was used since on visual inspection this appeared to optimally extract the differential bias field from the difference image. The repeat image is then corrected with the differential bias field, to give a new image pair with no differential bias field, but with the same absolute bias across both images.

Substituting these bias corrected images into Eq. 2 gives:

$$BBSI = \sum_{boundary} [K/(J_2 - J_1)[clip(v_b(x), J_2, J_1) - clip(v_r(x)b_1(x)/b_2(x), J_2, J_1)] \quad (4)$$

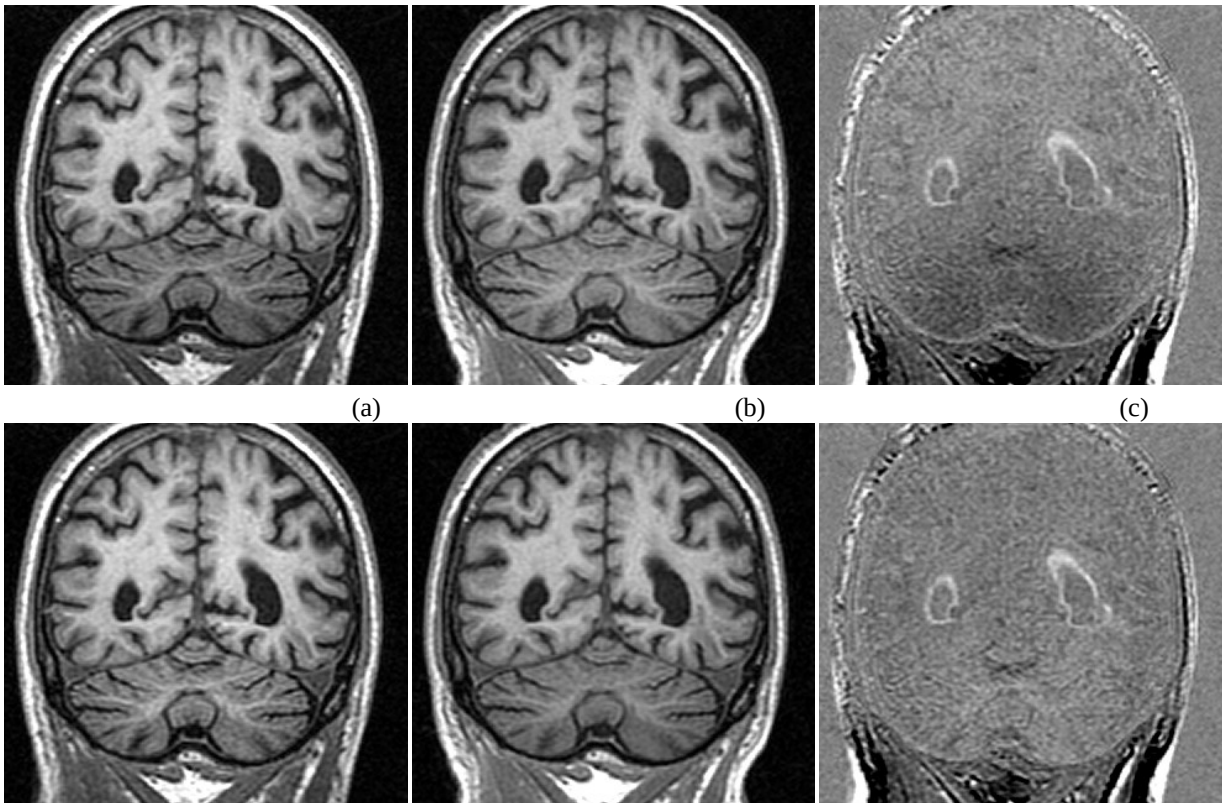
Eq. 4 is easily shown to be equivalent to calculating the BBSI using the true signal over a moving window given by limits $J_1 = I_1/b_1(x)$ and $J_2 = I_2/b_1(x)$. Given that the maximum bias is in the order of 20% , this window should lie within the boundary transition and hence give an accurate estimate of atrophy.

2.2 Application

A dataset (*full set*) of 280 longitudinal T1 weighted volumetric MR scan pairs (256x256 matrix, 124 slices, voxel size 1x1x1.5mm^3) were acquired of patients with Alzheimer's disease with a mean scan interval of one year. These longitudinal pairs were visually inspected by an expert for differential bias and a subset (*bias set*) of 19 pairs with significant differential bias were selected. The BBSI was calculated for each longitudinal scan pair. The differential bias correction technique (dbc) was then applied to each scan pair and the BBSI recalculated. The mean and standard deviation of the uncorrected and corrected BBSIs were then compared.

3 Results

Figures 2(a) and 2(b) show coronal slices from baseline and repeat images with significant differential bias. Figures 2(d) and 2(e) show the same coronal slices from baseline and repeat images corrected using dbc. Difference images between baseline and repeat images are given in Figures 2(c) and 2(f) for the original and dbc-corrected image pairs respectively. The pre and post dbc difference images show that dbc has removed the differential bias: note that the intensity gradient, particularly noticeable in the cerebellum, has been removed, while the real signal around the ventricles due to the ventricles expanding has been preserved.



(a) (b) (c)

(d) (d) (f)

Figure 2. Coronal slices through (a) baseline and (b) registered repeat image for longitudinal pair with significant differential bias; (d) and (e) show the same coronal slices after dbc; difference images between baseline and registered repeat image (c) and after dbc (f): note in (c) the artifactual difference in the cerebellum due to the differential bias and the real difference around the ventricles due to atrophy, while for the corrected pair in (f) the artifactual difference in the cerebellum has been removed by dbc while the genuine difference around the ventricles is unaffected.

Applying dbc significantly reduced the variance in the atrophy measurement of the bias set by $\sim 50\%$ ($p < 0.0001$) with a small reduction in the mean. This reduction in sd is also seen in the full set. The sd in the bias set is larger than in the full set because differential bias adds noise to the atrophy measurement.

	Bias set BBSI mean(standard deviation)	Full set BBSI mean(standard deviation)
non-dbc	13.7 (31.1)	20.1(19.4)
dbc	13.4 (20.5)	17.9(15.0)

Table 1. Mean and standard deviation of BBSI atrophy measure for full set (280 image pairs) and bias subset (19 image pairs) before and after dbc

4 Discussion

Longitudinal quantification techniques of serial MRI are increasingly important in neurological research and clinical trials. These measurements are significantly and materially impaired by differential bias. The technique presented here provides a non-parametric and non-iterative method for calculating differential intensity inhomogeneity. It makes no assumptions about tissue models or bias field models, which are generally necessary for the calculation of absolute bias. It assumes only that that the inhomogeneity field is relatively large scale in comparison to registration error, atrophy and noise.

Applying dbc to real longitudinal scans with significant differential bias removes differential bias between longitudinal scans thereby significantly reducing the variance in measurement of atrophy. Post-dbc atrophy measures therefore lead to reduced error in the assessment of longitudinal change, enabling better diagnosis, disease monitoring and measurement of therapeutic effects in serial MRI studies.

Future work includes further analysis of the spectral form of bias, reigstration error, atrophy and noise in order to select the optimum filtering technique for extracting the bias field. It would also be of interest to investigate the effect of dbc on the joint intensity histogram, which would be expected to be sharpened by the application of dbc. This technique could potentially be used to calculate absolute bias by registering a template image with zero bias to the target image and applying dbc.

References

1. P. A. Freeborough & N. C. Fox. "The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat mri." *IEEE Trans Med Im* **16**, pp. 623–629, 1997.
2. K. V. Narayana P. A., Brey W. W. & S. C. L. "Compensation for surface coil sensitivity variation in magnetic resonance imaging." *Magnetic Resonance Imaging* **6**, pp. 271–274, 1998.
3. S. R. & W. P. "Imaging of the active b1 field in vivo." *Magnetic Resonance in Medicine* **35**, pp. 246–251, 1996.
4. G. R. Tincher M., Meyer C. R. & W. D. M. "Polynomial modeling and reduction of rf body coil spatial inhomogeneity in mri." *IEEE* **12**, pp. 361–365, 1993.
5. S. G. Styner M., Brechbuhler C. & G. G. "Parametric estimate of intensity inhomogeneities." *IEEE* **19**, pp. 153–165, 2000.
6. Z. A. P. Sled J. G. & E. A. C. "A nonparametric method for automatic correction of intensity nonuniformity in mri data." *IEEE Trans Med Im* **17**, pp. 87–97, 1998.
7. B. P. Petrou M. *Image Processing: the fundamentals*. Wiley, 1999.
8. P. R. B. *Fundamental enhancement techniques. In Handbook of Medical Imaging: Processing and Analysis. Editor Bankman, I.N.* Academic Press, 2000.

Combining Topological and Geometric Features of Mammograms to Detect Masses

Byung-Woo Hong[1] Matthew Mellor[1] Stefano Soatto[2]
Michael Brady[1]

[1]University of Oxford, Oxford, U.K., [2]University of California Los Angeles, L.A., U.S.A.

Abstract. This paper presents a novel method for detecting masses in mammograms. Both topological (salience) and geometric (primarily texture) are used as features to characterise. Experimental results demonstrate that this combination of features is robust both for the segmentation and for the identification of masses.

1 Introduction

Breast cancer is the leading cause of death from cancer among women in many countries. Detecting a breast cancer at the earliest stage possible has the most important impact on prognosis. Mammography is the most cost effective method to detect early signs of breast cancer. However, mammograms are complex, textured images and there is substantial variation across the screening population, leading to difficulties in detection and diagnosis. In addition, the number of mammograms to be analysed in the screening programme is vast (eg 3 million annually in the UK alone). This is the motivation for the development of computer-aided diagnosis (CAD) systems that provide a consistent and reproducible second opinion to a radiologist. Unfortunately, progress toward building CAD systems for mass detection has been considerably slow, not least because of the subtle characteristics of their appearance [1]. While most previous work has been based on pixel-based statistical approaches, we aim to delineate regions corresponding to masses by analysing both the topological and geometrical structure of an image. Based, initially on regions of interest extracted by the algorithm presented in our previous work [2], we develop in this paper a two stage strategy. First, the mammogram is segmented into candidate mass regions based on the mammogram topology. This stage selects all salient regions that include masses along with a number of false regions that are candidates for masses. Then the segmented candidate mass regions are classified as masses or rejected on the basis of a measure of topological saliency combined with geometric texture features.

2 Segmentation

The first stage of our method is to segment the mammogram into salient regions which include all mass regions but also a small number of normal dense bright regions which constitute false positives. A segmentation method using topological properties of image structure has been presented and an efficient implementation is also provided.

2.1 Salient Region

Masses generally appear to be dense bright regions. The algorithm for detecting masses initially delineates the boundary of salient regions characterised by intensity blobs. The proposed segmentation algorithm analyses the topological structure of an image to detect salient regions as candidate masses. In analysing image structure, an image $I(\mathbf{x}), \mathbf{x} \in \Omega \subset \mathbb{R}^2$ is considered as a surface $\phi : \Omega \to \mathbb{R}^+$ where the intensity is regarded as the height of the surface. A region $R(t)$ and a level curve $\Gamma(t)$ at a level t on a surface ϕ is defined by:

$$\begin{aligned} R(t) &= \{\mathbf{x} | \phi(\mathbf{x}) > t\} \\ \Gamma(t) &= \partial R(t) \end{aligned}$$

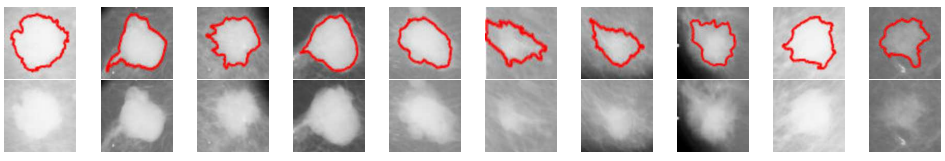
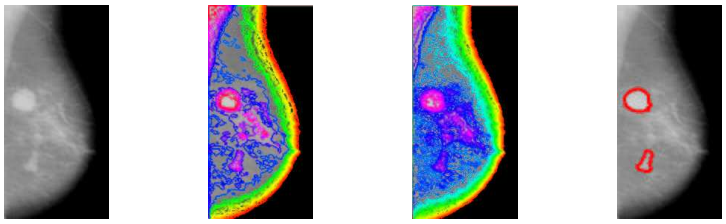


Figure 1. Examples for the segmentation of topologically salient regions. Top: Most salient base contours superimposed on the mammograms. Bottom: Original mammograms.



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

Figure 2. (a) A mammogram. (b) Coarse-scale contour map (N=30). (c) Fine-scale contour map (N=100). (d) Detected salient regions after discarding contours for the breast boundary and the pectoral muscle from (c).

A level curve is constrained to be a simple closed curve, thus the complement of a curve Γ consists of exactly two disjoint regions, an interior and an exterior. By definition of the level curve, a surface within an interior region of the level curve always forms a blob region that is higher than its local neighbourhood. In observing a topological change of a surface, saddle points and local maximum points are of interest. Let $\mathrm{H}(\phi)$ be the Hessian of a surface ϕ and $\det(\mathrm{H}(\phi))$ be its determinant. The set of all local maximum points on a surface ϕ within a region R is given by:

$$\mathbf{P}_M(R) = \{\mathbf{x} | \nabla\phi(\mathbf{x}) = 0, \det(\mathrm{H}(\phi)) > 0, \frac{\partial^2\phi}{\partial x^2} < 0, \frac{\partial^2\phi}{\partial y^2} < 0, \forall \mathbf{x} \in R\}$$

Similarly, the set of all saddle points on a surface ϕ within a region R is given by:

$$\mathbf{P}_S(R) = \{\mathbf{x} | \nabla\phi(\mathbf{x}) = 0, \det(\mathrm{H}(\phi)) < 0, \forall \mathbf{x} \in R\}$$

A saddle point implies a topological change of a surface, since the split of an object occurs at a saddle point as the level is elevated. For a given image $I(\mathbf{x})$, all saddle points $\mathbf{P}_S(\Omega)$ on a surface ϕ are initially determined for segmenting salient regions. Once the saddle points are known, all level curves at saddle points, called the *base level curves*, are given as:

$$\Gamma^* = \{\Gamma_i | \Gamma_i = \partial R_i(t), t = I(p), \forall p \in \mathbf{P}_S\}$$

Then, each base level curve $\Gamma_i \in \Gamma^*$ provides segmentation for the boundary of a candidate salient region R_i. The details of the segmentation algorithm are described in [2].

3 Region Classification

The list of mass candidate regions defined by the base level curves includes all real masses together with a number of false positives. In order to reduce the false positive rate of the system, features that are important for identifying masses are developed based on topological saliency measures and geometrical texture features. These features are used for classifying salient regions obtained by the segmentation into masses or normal dense tissues.

3.1 Topological Saliency Measures

Structural Contrast : A structure-based contrast measure is proposed for measuring the saliency of a region. The structural contrast of a region is defined as the difference between the intensity of a local maximum point within the region and the intensity of the base level curve that delineates that region. The minimum structural contrast $K(\Gamma(t))$ for a base level curve $\Gamma(t)$ on an image surface $\phi(\mathbf{x}, I(\mathbf{x}))$ is defined as:

$$K(\Gamma(t)) = \min_{\mathbf{x}_i \in \mathbf{P}_M(\Gamma(t))} (I(\mathbf{x}_i) - t)$$

If there is more than one local maximum within the region, the minimum contrast is defined so that it also takes into account the variance of the image structure.

Structural Variety : The structural variety implies a topological saliency in that a salient region is likely to be formed by one object that leads the region to have one local maximum. The structural variety $V(\Gamma(t))$ for a base level curve $\Gamma(t)$ is given by the number of local maximum points within $Int(\Gamma(t))$ as follows:

$$V(\Gamma(t)) = \frac{|\mathbf{P}_M(\Gamma(t))|}{A(\Gamma(t))}$$

where $A(\Gamma(t))$ is the area of the region within $\Gamma(t)$. The larger regions will naturally contain more local maximum points, so the measure for the structural variety is normalised by the area.

3.2 Geometric Texture Features

The texture features used are amplitude, contrast and orientation invariant local geometrical descriptions of the intensity surface. These provide a rich description of the image surface that complements the topological features. The features are derived from dimensionless combinations of linear filter responses. The basis for these filters is a rotationally symmetric filter $f(r)$. From this filter, further filters are generated by multiplying $f(r)$ by angular trigonometric functions:

$$\begin{aligned} h_n(r,\theta) &= f(r) * \cos(n\theta) \\ l_n(r,\theta) &= f(r) * \sin(n\theta) \end{aligned}$$

where $r = \|\mathbf{x}\|_2$ and $\theta = \tan^{-1}(x_1/x_2)$. These filters convolved individually with the image regions. A family of contrast invariant local symmetry descriptors can be generated from the responses to these filters as follows:

$$\Phi_n(\mathbf{x}) = \tan^{-1} \frac{(f \otimes I(\mathbf{x}))}{\sqrt{(h_n \otimes f \otimes I(\mathbf{x}))^2 + (l_n \otimes f \otimes I(\mathbf{x}))^2}}$$

These descriptors have values in the range $[0, \pi]$. Φ_1 is equivalent to the phase of the monogenic signal [3], and describes odd-even local symmetry. Φ_2 is related to the shape index [4] (if f_r is a Laplacian of a Gaussian, they are equivalent) and describes local even-even symmetry (i.e. ridges, troughs, saddle points and local maxima and minima). The local symmetry descriptors are each accompanied by an orientation. The offset between the orientation of each symmetry and the local principal orientation provides an additional local descriptor. The local orientations are encoded by phasors:

$$O_n(\mathbf{x}) = h_n(\mathbf{x}) + i \cdot l_n(\mathbf{x})$$

The angle of each phasor is a measure of the local orientation of each symmetry descriptor. ($\theta_n = \tan^{-1}(h_n/l_n)/n$). For the results presented in this paper, only the first three descriptors were used. The local orientations of each descriptor are normalised by dividing by a local orientation phasor:

$$\begin{aligned} O_m(\mathbf{x}) &= \frac{O_1(\mathbf{x})^3}{\|O_1\|^2} + \frac{O_2(\mathbf{x})^2}{\|O_2(\mathbf{x})\|} + O_3(\mathbf{x}) \\ O'_n(\mathbf{x}) &= \frac{O_n(\mathbf{x})\|O_m(\mathbf{x})\|}{O_m(\mathbf{x})\|O_n(\mathbf{x})\|} \end{aligned}$$

This gives six invariant descriptors, Φ_{1-3} and O'_{1-3}. The filters used here are scale and affine robust:

$$f(r) = \frac{A}{r^{\alpha+\beta}} - \frac{B}{r^{\alpha-\beta}}$$

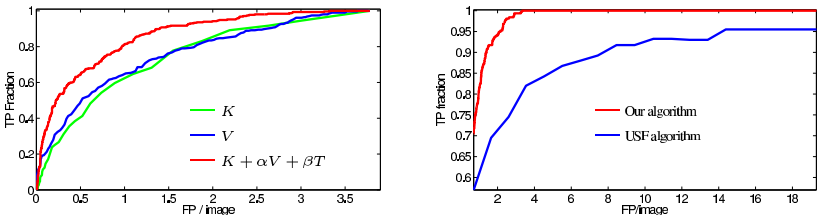
Two members of this family were used, giving twelve descriptors in all. The two filters are given by $\alpha = [2.75, 3.75]$, $\beta = 0.25$, with A and B chosen to give a zero DC component (that is, provide a measure that is invariant to brightness). Textures are characterised by the distribution of energy over these descriptors, and in particular the co-occurrence of energy in the different descriptors. It is impractical to learn the co-occurrence of all descriptors because of the high dimensionality. Instead, the joint energy distributions of each possible pair were estimated, giving a total of 66 two-dimensional distributions. The energy distributions are calculated as follows. In order to learn characteristic features for masses, the training set consists of a list of mass candidates, each manually labelled as belonging to one of two classes, true positive or false positive. For all mass candidates in a class, the six descriptors described above are calculated, along with the local energy (the square root of the sum of squares of all thirteen filter responses at each point). For each of the 66 possible descriptor pairings, a local energy weighted histogram is computed, the set of which acts as a region texture descriptor. The variance of each histogram location is also estimated. To classify, a variance weighted energy distribution difference is calculated:

$$E_{\text{diff}} = \frac{E_{FP} - E_{TP}}{\sqrt{V_1 + V_2 + \frac{1}{M}}}$$

where M is the number of histogram bins, here 50x50 histograms were used (N=2500). The $1/M$ term is added to reduce the influence of parts of the histogram which have had very few votes. For each unclassified mass candidate the energy distribution, $E(\Gamma(t))$ is computed as above. The region is then given a score, $T(\Gamma(t))$, by taking the inner product of its energy distribution with E_{diff}:

$$T(\Gamma(t)) = \sum E_n * E_{\text{diff}}$$

$T(\Gamma(t))$ may be treated as a "texture saliency" measure, to complement the topological saliency measures.



(a) (b)

Figure 3. (a) ROC operated by different saliency measures. (b) ROC comparison with the USF algorithm.

4 Experiments

We present experimental results based on a set of 400 mammograms with masses varying in size and subtlety. Mammograms were selected from various pathological categories in the USF database [5]. The resolution of the data set was 200μm per pixel. The detection of masses was determined by the ratio of the intersection between a segmented region and the ground truth over the union of them. In the segmentation stage, a coarse-scale contour map ($N = 30$) was initially used to localise topologically salient regions characterised by base contours. Among the salient regions, the breast boundary and the pectoral muscle were excluded by using their strong anatomical constraints in terms of size and location. Then, the salient contours that correspond to the ones in the coarse-scale contour map were identified in a fine-scale contour map ($N = 100$) by traversing inclusion tree. The topological saliencies were measured based on the fine-scale contour map and the geometric texture descriptor was measured based on the interior regions of salient contours selected from the fine-scale contour map. More accurate both segmentation and saliency measurement were obtained from the fine-scale contour map. An example for contour maps with different scales and extracted base contours are shown in Figure 2. The base contours with high minimum structural contrast and low structural variety define the boundaries of salient regions for mass candidates. The algorithm achieved 100% detection rate with 3.3 false positives per image and 90% with 1.4 respectively. The performance of the algorithm was evaluated by FROC operated by the saliency measure, K, V, and $K + \alpha V + \beta T$ with the optimal choice of weights as presented in Figure 3(a). We have used no priori information about mass at measuring K and V, but we have trained textures of mass using 100 samples. An additional texture information to topological saliency improved the results as shown in Figure 3(a). The comparison of our results with the USF algorithm [5] is presented in Figure 3(b). In addition to the improved performance, our algorithm provides exact boundary of masses as opposed to just finding a point within the mass region.

5 Conclusion

We have developed a blend of topological and geometrical descriptors that together effectively characterise mass features. A global structural approach based on a topographic representation is shown to be a useful counterpart to a local statistical approach for detecting masses in mammograms. Structural saliency measures are shown to be useful mammographic features to delineate regions of interest and the texture information using phase information appears to be an informative characteristic feature of mass. Experimental results indicate that this method can be used as the basis for an effective prompting tool to assist radiologist in the diagnosis of breast cancer.

References

1. R. Highnam & J. Brady. *Mammographic Image Analysis*. Kluwer Academic, 1999.
2. B.-W. Hong & J. Brady. "A topographic representation for mammogram segmentation." *In MICCAI '03* **32**, pp. 675–695, 2003.
3. M. Felsberg & G. Sommer. "Image features based on a new approach to 2d rotation invariant quadrature filters." In *ECCV'02*, pp. 369–383. 2002.
4. J. Koenderink & A. van Doorn. "Surface shape and curvature scales." *IVC* **10**, pp. 557–565, 1992.
5. M. Heath, K. Bowyer, D. Kopans et al. "The digital database for screening mammography." *In IWDM'00* 2000.

Tracking of myocardial walls and study of contractility and thickening

Ramn Casero-Cañas[a], Jonathan Timperley[b], J. Alison Noble[a] and Harald Becher[b]

[a]Medical Vision Laboratory, University of Oxford, United Kingdom
[b]John Radcliffe Hospital, Oxford, United Kingdom

Abstract. The study of cardiac muscle contractility is fundamental in the evaluation of cardiac function. We propose a continuous measure of contractility based on segment thickening in echocardiographic contrast B-mode time series. Thickness is computed with the symmetric nearest neighbour distance method from the inner to the outer myocardial walls. Preliminary results from 2 studies are compared to nominal scoring by an expert and illustrate how this measure can be useful in the detection of hypokinetic segments.

1 Introduction

Contractility is a term that in cardiology refers to the ability of the cardiac muscle (myocardium) to shrink and stretch to pump blood with the right pressure and timing. Multiple cardiac pathologies produce abnormal regional contractility that, if detected, is of critical importance for diagnosis and may enable preemptive treatment. Of special interest are non-invasive methods, amongst which transthoracic echocardiography (or standard echo) is one of the most widely used because of its safety, speed, low cost and lack of radiation. Transthoracic echocardiography is a medical data acquisition technique based on ultrasound waves that are sent from a transducer placed on the chest so that they intersect the heart. By processing the sound wave echoes caused by acoustic impedance mismatches it is possible to obtain structural, motion or velocity information of heart tissues and fluids. In the case of standard echo, the left ventricle (LV) is favoured over other cavities in research, not only because the oxygen-rich blood is pumped out from it at high pressure into the arteries and is the part of the heart most likely to have a problem, but also because it is most conveniently positioned near the probe. There is not a unique quantitative measure for contractility, so several indirect parameters are targeted for its assessment, for example the Ejection Fraction (EF, the fraction of end-diastolic volume ejected in systole by the LV), inner wall (endocardial) wall motion i.e. excursion or velocity, myocardium thickening and strain. See e.g. Garca-Fernndez et al. [1] for a review. Other authors work with intraventricular preassure gradients [2] or tissue texture classification [3], although the latter requires invasive catheterization for the probe.

Contractility problems appear generally in local neighbourhoods of the myocardium. Global parameters such as the EF are unsuitable for detecting myocardial defects, as the heart can compensate for a local damaged region by making the rest work harder so that the global function is not affected. Such a measure can also not locate where the problem is. Other parameters such as wall motion, thickening and strain are implicitily local, but they are generally averaged for a given myocardial segment, rather than evaluated in the muscle as a continuous function. There are segment models with 9 to 20 segments, although the commonest in echocardiography is the 16-segment model [4]. In clinical practice, assessment of segment contractility is performed with the nominal method of scoring, where each segment is classified into one of the following categories: normokinetic (healthy), hypokinetic (reduced contractility), akinetic (dead) and dyskinetic (when there is a paradoxic outward bulging during systole). Thus clinicians use an evaluation function f, such that $f : \Omega \mapsto \Phi$ where Ω is the discrete domain of segments and Φ is the discrete range of scoring. Jacob et al. [5] proposed functions f for the evaluation of endocardial excursion and myocardial thickening as $f : Q \mapsto \Re$, i.e. the parameters are real variables evaluated over the discrete domain of B-spline control points Q. Contours are represented as uniform parametric B-splines with N control points for both the endocardium $\{Q_{\text{en},i}\}_{i=1}^{N}$ and epicardium $\{Q_{\text{ep},i}\}_{i=1}^{N}$, where Q_i are the cartesian coordinates of the i-th control point. Thickening for the two i-th corresponding control points $Q_{\text{Diff},i}$ is computed as the difference $Q_{\text{Diff},i} = Q_{\text{ep},i} - Q_{\text{en},i}$. Thickening $Q_{\text{Diff},\omega}$ for a given segment $\omega \in \Omega$ is computed as $Q_{\text{Diff},\omega} = 1/J \sum_j Q_{\text{Diff},j}$, such that there are J control points $Q_{\text{Diff},j} \in \omega$. However, this approach has some drawbacks: (1) Thickening for a segment would be better defined as $Q_{\text{Diff},\omega} = 1/J \sum_j \|Q_{\text{Diff},j}\|_2$. (2) We are limited to interpolating splines, i.e. control points are points of the curve, because only in this case is $Q_{\text{Diff},i}$ meaningful. (3) To increase the number of points for which thickening is measured, one has to increase the number of control points, and recompute the training set and the Principal Component Analysis for the new control points.

In this paper we propose a thickening measure to evaluate contractility based on the symmetric nearest neighbour method [6]. The details are provided in section 2. We have tested this measure with 2 B-mode contrast echo studies, where each study represents a complete cardiac cycle. B-mode is an echocardiograpy modality where grey-scale time series of a two-dimensional plane across the heart are acquired. It is possible to enhance the signal-to-noise ratio of the B-mode image injecting a gas-filled microbubble suspension (called a contrast agent) into the patient's blood stream. The safety and tolerance of contrast echocardiography are extremely favourable and the technique is considered to be cost effective [7]. The inner and outer wall of the myocardium were segmented automatically with a prototype commercial tracker called Quamus®. This paper reports for the first time on how this system can be used for automatic epicardial as well as endocardial border tracking.

2 Contractility measure based on thickening

The tracking algorithm represents contours as interpolating uniform B-splines. These contours are sampled uniformly to obtain a vector of coordinates $p = (p_1, p_2, \dots p_N)$ for the the inner contour (endocardium) and a vector $r = (r_1, r_2, \dots r_M)$ for the outer contour (commonly called epicardium, although e.g. in 4C the outer wall is part LV epicardium and part right ventricle endocardium). We propose to use the distance measure obtained from applying the symmetric nearest neighbour (SNN) method [6] to the myocardial inner and outer wall contours to compute myocardial thickening as $T = T(p)$ such that

$$T(p_i) = \|p_i - r_j\|_2 \tag{1}$$

where p_i and r_j are correspondent points. The reason to use p rather than r is that in B-mode time series the uncertainty in the segmentation of the inner wall is lower than in the outer wall. Thus clinicians work with a domain of segments Ω (typical dimension for the 4C plane is 6 elements), Jacob et al. [5] work with a domain of control points Q (typical dimension is 24), and our approach works with a domain of any sampling p of the endocardial contour (typical dimension is 300). The SNN method finds a correspondence between p and r. First, for each point p_i in the endocardium, the nearest neighbour r_j in the epicardium is obtained. Then, if $p_l = p_i$, where p_l is the nearest neighbour of r_j in the endocardium, we say that p_i and r_j are symmetric nearest neighbours and the match is kept. In a second run, a correspondence is computed for the sets of K consecutive unmatched points $p_{i+1}, \dots, p_{i+K}$, drawing equally spaced points in the epicardium, between the points r_j and r_{j+K+1}. In the two test studies, we have got between 20% and 30% of symmetric nearest neighbour points from a total of $N = M = 300$ points for each contour in every frame. Figure 1 shows an example of the distance measure between the inner and outer walls of the myocardium as seen in a 4C view. In this figure we can see a line that links every $p_i \in p$ to its correspondent point $r_j \in r$. $T(p_i)$ is the length of that line. Finally, we propose to evaluate contractility C as the ratio between maximum thickening $T_{\max}$ and thickening in end-diastole T_{ED}.

$$C(p) = \frac{T_{\max}(p)}{T_{\text{ED}}(p)} \tag{2}$$

3 Experimentation and Results

Quamus® has previously been validated against expert hand tracing and cine-MRI [8] for endocardial tracking. We have used it to track the epicardium in the 2 test studies too, and to assess its performance we have compared automatic tracking to a single expert hand tracing in 3 planes: 2 chamber (2C), 3 chamber (3C) and 4C. The SNN method described in the previous section was used to find the distances between contours. The results are summarized in table 1. The variability for automatic tracking with 2 different initializations of the tracker was found to be between 0.79 mm and 1.05 mm. The variability for a human expert drawing the same epicardial contours 2 weeks later (intraobserver variability) is between 0.94 mm and 1.32 mm. Thus the automatic tracker is more consistent in its results than the human expert. The variability between hand traced contours, that we use as our pseudo-gold standard, and automatic tracked contours, is between 1.18 mm and 1.46 mm, a result that in the worst case is only a 10.6% bigger than intraobserver variability. There is, to our best knowledge, no continuous real measure of contractility based on thickening against which we can validate the one that we propose in Eqn. 2. To illustrate how the proposed measure can be used for diagnosis, we divided the myocardium in the 4C plane into 6 segments following the 16-segment model [4]. For each segment ω we average the thickening of all the endocardial points $p_i \in \omega$ such that $\bar{T}(\omega) = 1/L_\omega \sum_{L_\omega} T(p_i)$ where L_ω is the number of endocardial points

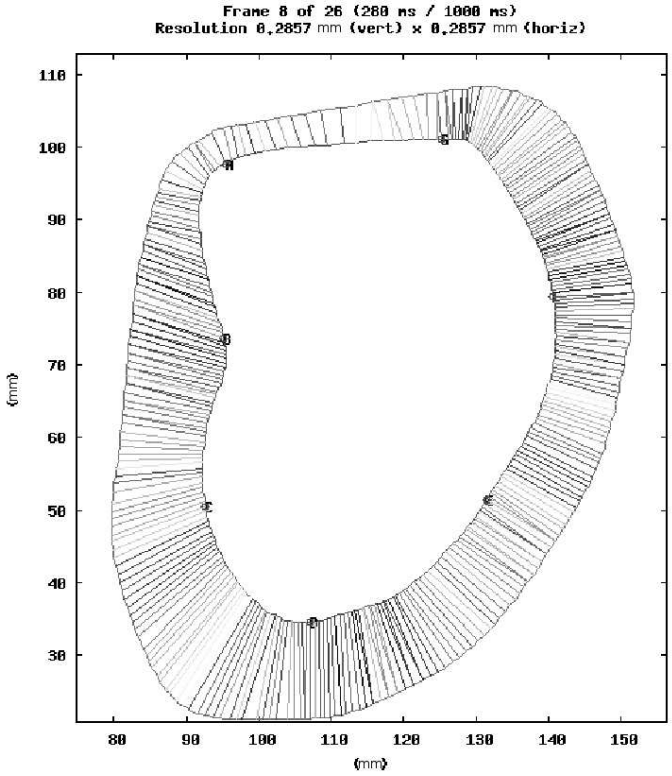


Figure 1. Distance between two contours computed with the symmetric nearest neighbour method for a frame of a 4C view. Lateral wall is on the right, septum is on the left, apex is at the bottom and the base is on top.

that belong in segment ω. Then contractility $C(\omega)$ was computed for each segment as in Eqn. 2. This value is compared to the scoring value assigned to that segment by an expert observer in table 2. The ratio (study 1)/(study 2) between $C(\omega)$ values for segments BS, MS and AS in the two studies is between 0.8 and 0.9, which shows good correlation, whereas for segments AL, ML and BL, that in the first study are normokinetic but in the second are hypokinetic, is between 0.4 and 0.5.

Plane	A-A (mm)	H-H (mm)	A-H (mm)
2C	1.05	1.32	1.46
3C	0.79	1.13	1.28
4C	0.87	0.94	1.18

Table 1. Automatic epicardial tracking compared to expert hand tracing. A-A: Variability between automatically traced contours. H-H: Intraobserver variability. A-H: Variability between automatic and hand traced contours.

4 Conclusions

A quantitative measure for contractility based on myocardial thickening over a domain of any sampling of the endocardial contour has been proposed. Automatic tracking of the epicardium is only 10.6% worse and has better reproducibility than expert hand tracing, according to some preliminary results, although it remains to perform an extensive validation. Finally, it has been illustrated how the proposed measure is consistent with evaluation by scoring in normokinetic and hypokinetic segments. The obvious next steps are to define the contractility function $C : [0, 1] \mapsto \Re$, i.e. to define contractility over a normalized continuous parametrization of the endocardium rather than over a sampling, and to perform extensive validation in a larger clinical study with comparision with a gold standard.

	BS	MS	AS	AL	ML	BL
study 1, eqn. 2 (%)	108	157	171	117	110	131
study 1, scoring	N	N	N	H	H	H
study 2, eqn. 2 (%)	136	175	186	225	272	265
study 2, scoring	N	N	N	N	N	N

Table 2. Comparision of contractility measure of eqn. 2 to expert scoring in 4C plane. N: normokinetic. H: hypokinetic. The segments are: BS, MS, AS (basal, mid and apical septum) and AL, ML, BL (apical, mid and basal lateral).

Acknowledgements

The work presented in this paper is funded by Mirada Solutions Ltd. and the Engineering and Physical Sciences Research Council (EPSRC). Quamus® is a software application developed on an European Community project involving Mirada, the Hospital Universitario Gregorio Marañón of Madrid, the Medical Vision Laboratory of the University of Oxford and the John Radcliffe Hospital of Oxford.

References

1. M. . Garca-Fernndez, J. Bermejo, E. Prez-David et al. "New techniques for the assessment of regional left ventricular wall motion." *Echocardiography* **20(7)**, pp. 659–672, 2003.
2. R. Yotti, J. Bermejo, J. C. Antoranz et al. "Noninvasive assessment of ejection intraventricular pressure gradients." *J. Am. Coll. Card.* **43(9)**, pp. 1654–1662, May 2004.
3. X. Hao, C. J. Bruce, C. Pislaru et al. "Segmenting high-frequency intracardiac ultrasound images of myocardium into infarcted, ischemic, and normal regions." *IEEE Trans. Med. Imaging* **20(12)**, pp. 1373–1383, December 2001.
4. M. D. Cerqueira, N. J. Weissman, V. Dilsizian et al. "Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart." *Circulation* **105(4)**, pp. 539–542, 2002.
5. G. Jacob, J. A. Noble, C. Behrenbruch et al. "A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography." *IEEE Trans. Med. Imaging* **21(3)**, pp. 226–238, 2002.
6. X. Papademetris. *Estimation of 3D Left Ventricular Deformation from Medical Images Using Biomechanical Models*. Ph.D. thesis, Yale University, May 2000.
7. J. Timperley, H. Thibault & H. Becher. *The Oxford Way of Doing Contrast Echocardiography*, chapter 14, pp. 235–249. In [9], 2004.
8. ADEQUATE Consortium Final Report. "[Journal article in preparation].", 2004.
9. J. L. Zamorano & M. . Garca-Fernndez. *Contrast Echocardiography in Clinical Practice*, volume XIII. Springer Verlag, 2004.

Towards Classification of Prostate MRI

Reyer Zwiggelaar[a]*, Yanong Zhu[b] and Stuart Williams[c]

[a]Department of Computer Science, University or Wales, Aberystwyth
[b]School of Computing Sciences, University of East Anglia, Norwich
[c]Department of Radiology, Norfolk and Norwich University Hospital, Norwich

Abstract. This study investigates if grey-level profiles orthogonal to the prostate boundary contain enough information to establish if prostate cancer is confined to the gland or whether it has spread into the periprostatic tissues. This thesis is based on the notion that profiles can be radically changed at locations where a cancer extends out of the gland. Statistical modelling, to generalise the profiles and reduce noise, in combination with nearest neighbour classification is used to classify individual profiles. Sensitivity and specificity measures are provided based on a leave-one-patient-out methodology. Limitations of the presented approach and available data are discussed.

1 Introduction

Prostate cancer is now the most frequently diagnosed male malignancy, with one in every 11 men developing the disease [1]. It is the second most common cause of cancer deaths in men. Patients who present with organ confined disease may be suitable for surgery - radical prostatectomy and bilateral pelvic lymph node dissection. However, once tumour has spread beyond the gland, radical radiotherapy is the preferred option. In advanced disease, hormone deprivation, radiotherapy and chemotherapy all have a role in patient management. The importance of imaging is to determine whether tumour is confined to the gland or whether tumour has spread into the periprostatic tissues. The detection of nodal and bone marrow spread is also important. Magnetic Resonance Imaging (MRI) is now the staging method of choice for cases of proven prostate cancer [2,3]. It is the most reliable technique for the depiction of the zonal anatomy of the prostate - 70% of tumour arising from the peripheral zone of the gland. Its superior contrast resolution and multiplanar capabilities allow the best chance of detecting extracapsular extension of tumour. Nevertheless, early periprostatic spread can be subtle, with intra-observer discrepancies noted.

The main aim of the developed approach is to improve the assessment of the spread of cancer both within and outside of the gland. MRI provides three-dimensional anatomical information displayed as two-dimensional slices. The overall aim of the project is to investigate the information contained within a number of grey-level profiles which are extracted orthogonal to the boundary of the prostate. The profiles for normal prostates are characteristic showing a number of transitions between anatomical features within the gland. When a cancer extends out of the gland, these profiles can be radically changed. To be able to extract these profiles the prostate needs to be manually or automatically segmented. We present initial classification results.

It should be made clear that the presented work is based on manual segmentation of the prostate by expert radiologists. However, related work has investigated the automatic segmentation of prostate from MRI data. Two distinct approaches have been developed which both show a strong correlation with expert annotations [4]. Work based on a polar transform approach used a few basic assumptions about the prostate and the available image data [5]. In addition, standard 2D and 3D ASM have been investigated [4]. Subsequently, a hybrid 2D+3D ASM approach has been developed which showed significant improvement over 2D and 3D ASM modelling [6]. For this hybrid approach the mean root-mean-square-distance from the segmentation results to the manual annotations of the prostate boundary was 5.4, with a standard deviation equal to 2.9. A more detailed discussion of the hybrid 2D+3D shape modelling approach can be found in an accompanying paper in these same proceedings [6].

Some of the presented work is closely related to the grey-level profile modelling used for the classification of linear structures in mammographic images [7,8]. The main difference is that in the mammographic case the grey-level profiles represented a range of tube-like structures whilst for the prostate the grey-level profiles represent the transition from the prostate to its surrounding tissue. In this initial investigation it is assumed that there are only two classes for the transition, either normal (cancer, if present is confined to the prostate) or abnormal (cancer has spread to the periprostatic tissue). To our knowledge, this is a unique approach to the classification and staging of prostate cancer using MRI data.

* email: rrz@aber.ac.uk. The presented research is supported by Prostate Research Campaign UK

2 Data

The main data consists of 20 prostate MRI volumes. All images are obtained on a 1.5 Tesla magnet (Signa, GE Medical Systems, Milwaukee, USA) using a phased array pelvic coil. Field of view 24×24 cm, matrix 256×512, slice thickness 3mm with an interslice gap of 0.5mm, TR 7800ms, TE 102ms. Fig. 1 shows a typical example from the data set. The prostate can be found in the centre of the image. There are minor benign hypertrophic changes in the central zone. The peripheral zone architecture is generally preserved, with some patchy loss of the normal high T2 signal, in keeping with some malignant infiltration. There is no extracapsular extension present. The shown slice and its volume are used as an examplar, but it should be noted that similar results have been obtained for the other volumes in the dataset.

All images were manually annotated by an expert radiologist. For each prostate a finite number of profiles are extracted. Here we have extracted a single profile for all the positions on the annotated prostate boundary. For a typical prostate this approach results in about two-hundred intensity profiles per slice.



Figure 1. Axial view prostate MRI example.

3 Methods

Once the prostate has been segmented it is possible to extract intensity profiles orthogonal to the boundary. The orientation of the extracted profiles is taken with respect to the local orientation of the segmented boundary. To obtain the local orientation, ϕ, at the boundary we have used linear regression given by

$$\phi = \frac{N \sum xy - \sum x \sum y}{N \sum x^2 - \sum x \sum x} \tag{1}$$

which is based on all N points of the segmented boundary found within a region of interest, where (x, y) are the coordinates of the boundary points and the summations are over all N points. The intensity profile is extracted orthogonal to ϕ with the centre of the profile the point on the segmented boundary. For all the profiles a set number of points on either side of the segmented boundary are extracted. Here we have used sixteen positions on either side of the boundary which makes the total length of the intensity profiles equal to thirty-three. We have used bi-linear interpolation to determine the intensity at the profile points.

From a classification point of view we have considered two options. The first is to use the intensity profiles as the feature vector for the classifier. A second approach uses principal component analysis [7, 9] to generalise the profiles, remove noise and reduce their dimensionality. The first n (we used n equal to eight in the experiments below) principal components are used as the feature vector for the classifier. We have used a leave-one-patient-out methodology in combination with a nearest neighbour approach to classify the profiles as being normal or abnormal. The nearest neighbour classifier uses a simple Euclidean distance metric. The regions of the prostate that contain abnormal areas in the slices are used as ground truth which have been marked by an expert radiologist.

4 Results

For each segmented prostate a finite number of profiles are extracted. As discussed in Sec. 3 for a typical prostate as displayed in Fig. 1, this approach results in about two-hundred intensity profiles per slice. A few typical examples of such intensity profiles can be found in Fig. 2. On the left this shows a profile across an abnormal boundary and on the right it shows a typical profile across a normal boundary. The presented profiles are typical in that the normal boundary example shows a distinct dip in the centre which represents the boundary between the prostate and the periprostatic tissue. Due to the presence of prostate cancer this boundary is blurred in the abnormal example.

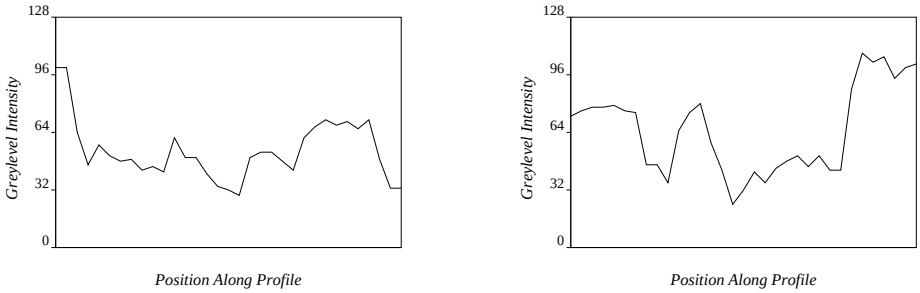


Figure 2. Example intensity profiles, where on the left abnormality and on the right normality is represented.

Using a leave-one-patient-out methodology, the results of a nearest neighbour abnormal profile classifier for an example prostate MRI slice can be found in Fig. 3. It should be clear that this shows strong correlation with the expert annotations for the abnormal profiles, but at the same time this goes at a high false positive cost. For all the profiles from the data the overall sensitivity is 63.6% whilst the specificity is 62.9%.

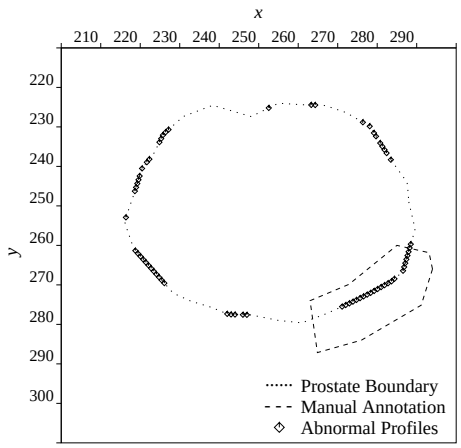


Figure 3. Abnormal profile classification results comparing a nearest neighbour classifier ($\diamond$) and manual annotation (dashed line). The results are based on a leave-one-patient-out methodology.

The equivalent leave-all-in results are respectively 81.8% and 99.2%. These results are based on all the available data (so the annotation for the slice to be classified is included in the training data). This shows good correlation between the nearest neighbour classifier and the expert annotations. These results indicate a potential improvement when the size of the data-set is increased, but at the same time indicates that data normalisation issues need further investigation.

The above results are based on using the first eight principal components to form the feature vector space for the nearest neighbour classifier. The number of principal components can be increased, providing a closer approximation of the original grey-level profiles. When increasing the number of used principal components to twenty-four the sensitivity and specificity for the leave-one-patient-out experiments are 58.4% and 60.3%, respectively. The equivalent leave-all-in results are 97.4% and 99.8%. This shows that data reduction is beneficial for the leave-one-patient-out approach, whilst the opposite is true for the leave-all-in experiments. This aspect needs further investigation with a first emphasis on increasing the number of annotated datasets.

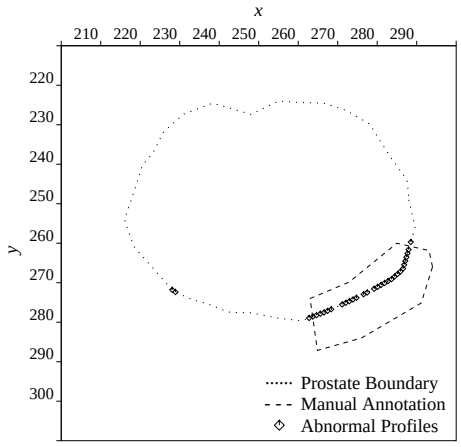


Figure 4. Abnormal profile classification results comparing a nearest neighbour classifier (◇) and manual annotation (dashed line). The results are based on including all the data in the modelling.

5 Discussion and Conclusions

When compared to modelling based on all the data, the results based on a leave-one-patient-out approach showed a degradation of the correlation between the classifier and the expert annotations. When modelling based on all the data, most of the true positive profiles rely on the profiles next to it to be classified correctly. The decrease in performance for the leave-one-patient-out evaluation might have been caused by the relative small size of the training data and this is an area of future development.

An additional area of investigation will be the normalisation of the grey-level profiles. The current lack of normalisation might be one of the causes for the limited performance of the leave-one-patient-out results.

A third area of improvement will be to take the correlation between adjacent grey-level profiles and slices into account. It is likely that such an approach will reduce the number of false positive detected areas. At the same time, the modelling might be extended to use grey-level profiles orthogonal to the 3D prostate shape instead of the 2D in-slice profiles used for the presented results.

In summary, profiles were extracted orthogonal to the segmented boundary. The profiles, in combination with statistical modelling, can be used in a nearest neighbour classifier to produce plausible results. The classification results show good correlation with expert annotations. Current research concentrates on improving these classification results.

References

1. S. Hayward & G. Cunha. "The prostate: Development and physiology." *Radiologic Clinics of North America* **38(1)**, pp. 1–15, 2000.
2. K. Yu & H. Hricak. "Imaging prostate cancer." *Radiologic Clinics of North America* **38(1)**, pp. 59–87, 2000.
3. M. Engelbrecht, G. Jager, R. Laheij et al. "Local staging of prostate cancer using magnetic resonance imaging: a meta-analysis." *European Radiology* **12(9)**, pp. 2294–2302, 2002.
4. Y. Zhu, R. Zwiggelaar & S. Williams. "Prostate segmentation: a comparative study." 7^{th} *Conference on Medical Image Understanding and Analysis* **Sheffield, UK**, pp. 129–132, 2003.
5. R. Zwiggelaar, L. Blot, D. Raba et al. "Set-permutation-occurrence matrix based texture segmentation." *Lecture Notes in Computer Science* **2652**, pp. 1099–1107, 2003.
6. Y. Zhu, S. Williams & R. Zwiggelaar. "Segmentation of volumetric prostate MRI data using hybrid 2D+3D shape modelling." 8^{th} *Conference on Medical Image Understanding and Analysis* **London, UK**, pp. this volume, 2004.
7. R. Zwiggelaar & C. Boggis. "Automatic classification of linear structures in mammographic images - is it possible?" 3^{rd} *Conference on Medical Image Understanding and Analysis* **Oxford, UK**, pp. 185–188, 1999.
8. R. Zwiggelaar, C. Taylor & C. Boggis. "Linear structures in mammographic images: Detection and classification." *IEEE Transactions on Medical Imaging* **in press**, 2004.
9. I. Jolliffe. *Principal Component Analysis*. Springer Verlag, 1986.

Partitioning the Cingulate Gyrus using Bézier curves

Richard G Boyes[a]*, Josephine Barnes[a] and Nick C Fox[ab]

[a]Dementia Research Centre, Institute of Neurology, Queen Square, London
[b]Department of Clinical Neurology, National Hospital for Neurology and Neurosurgery, Queen Square, London

1 Introduction

The cingulate gyrus is an arched convolution of grey matter that lies above the corpus collosum in the brain. A segmentation of the cingulate from a T1-weighted volumetric MRI scan can be seen in Figure 1.

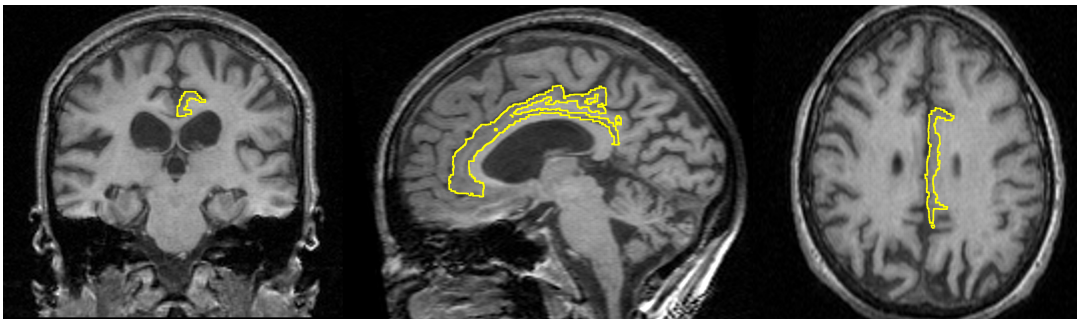


Figure 1. A segmented cingulate in coronal, sagittal and axial views

There is some evidence to suggest that pathological processes such as Alzheimer's disease (AD) affect the cingulate at an early stage. Moreover some sections may be affected differently from others [1]. Measuring atrophy in sections of the cingulate may therefore be a more accurate indicator of the disease as well as aiding in the study of the natural history of AD. In order to measure tissue atrophy of the cingulate in certain sections i.e., rostral anterior cingulate (RAC), caudal anterior cingulate (CAC) and posterior cingulate(PC) it is necessary to partition the cingulate in an accurate and reproducible way.

Typically the cingulate is segmented on T1-weighted volumetric MRI scans to define a binary mask that identifies the structure on the scan. This structure has convoluted folds that make it highly complex, with a large degree of inter-individual variation in gyral and sulcal patterns, making consistent labelling and subdivision difficult. Because the structure has folds it is difficult to reduce the structure to a simpler geometry by the use of morphological methods e.g., skeletonization. Previous methods have attempted to partition the cingulate manually, to our knowledge none have attempted to automatically partition the cingulate based on a previously defined mask. However, some methods have used a previously defined mask to partition other structures, such as the corpus callosum. These were based on creating and then partitioning an enclosing structure. [2].

Another method would be to find an approximation to the central axis of the cingulate and then partition the mask using a perpendicular plane to this axis. We have developed a method based on this concept. This method takes a cingulate region-of-interest (ROI) defined on a volumetric MR scan and fits (in a least squares sense) a cubic Bézier curve to this mask. This curve is used as a smooth approximation to the central axis of the cingulate, enabling the estimation of the length of the structure, and also the partitioning of the three dimensional mask into sections.

2 Methods

2.1 Theory

A one-dimensional cubic Bézier equation to parametrically interpolate a vector field takes the form:

$$\vec{x}(t) = (1-t)^3\vec{x}_1 + 3t(1-t)^2\vec{x}_2 + \ 3t^2(1-t)\vec{x}_3 + (t)^3\vec{x}_4 = \sum_{n=1}^{4} \psi_n(t)\vec{x}_n \qquad (1)$$

where $\vec{x}_n$ are control points in 3-d space and t is a local parametric coordinate.

*rboyes@dementia.ion.ucl.ac.uk

To fit Equation 1 to the cingulate mask, an iterative procedure is required. Firstly each point in the mask is orthogonally projected onto an estimate of the curve to obtain a value of t for each mask point, and then a function that minimises the squared distance between the curve and data points with respect to the control points is solved. The mask points are reprojected and the minimisation reapplied. Thus at each step we seek to minimise

$$F(x) = \sum_{d=1}^{D} (\vec{x}(t_d) - \vec{x}_d)^2 \tag{2}$$

where t_d is the parametric coordinate of the orthogonal projection of each mask point $\vec{x}_d$ onto the curve. To calculate t_d, a non-linear iterative procedure is required. Given a starting position t_s for the mask point projection, the geometric position of this projection is given by the interpolation formula in Equation 1. An error function can then be established as the Euclidean distance between the interpolated and the actual position of the data point. The position t that minimises this function can then be found using a root finding method such as Brent or Newton-Raphson on the derivative of this function. This t position is effectively the orthogonal projection of the data point onto the curve.

Substituting Equation 1 into 2,

$$F(x) = \sum_{d=1}^{D} (\sum_{n=1}^{4} \psi_n(t_d)\vec{x}_n - \vec{x}_d)^2 \tag{3}$$

To minimise this distance, we differentiate with respect to the curve parameters $\vec{x}_m$ and set the resulting equation to be zero:

$$2\sum_{d=1}^{D} (\sum_{n=1}^{4} \psi_n(t_d)\vec{x}_n - \vec{x}_d) \sum_{m=1}^{4} \psi_m(t_d) = 0 \tag{4}$$

or

$$\sum_{d=1}^{D} (\sum_{m=1}^{4} \psi_m(t_d) \sum_{n=1}^{4} \psi_n(t_d))\vec{x}_n = \sum_{d=1}^{D} (\sum_{m=1}^{4} \psi_m(t_d)\vec{x}_d) \tag{5}$$

Which represents a 4×4 system of equations $A_{mn}\vec{x}_n = b_m$. This is solved to find the nodal parameters $\vec{x}_n$ following a projection of the mask points onto the current curve. The projection is then updated and the minimisation re-applied. This process continues until the nodal parameters $\vec{x}_n$ have converged.

In practice two of nodal parameters $\vec{x}_1$ and $\vec{x}_4$ are manually fixed to be the observed start and end points of the cingulate structure.

Upon completion of the fit, each point $\vec{x}_d$ in the mask is then orthogonally projected onto the finished curve, and its parametric coordinate t_d is calculated. Thus if we wish to partition the cingulate mask into thirds, say, we would take all mask points with parametric coordinates in each of those thirds and create new masks from them.

2.2 Application

As a test for the consistency of the Bézier curve fitting method, we used T1 weighted scans of 11 healthy subjects at two time points, assuming no disease-related change in the cingulate. Both the whole cingulate and PC were segmented by an expert using MIDAS software [3] for each subject, and then the Bézier curve fitting method was applied to cleave the posterior third of the whole cingulate to obtain another PC mask. The proportion of the PC volume to the whole cingulate volume was calculated for all subjects using both methods. These proportions were

Atrophy %/year	Cingulate	RAC	CAC	PC
in controls	-0.68 ± 1.88	0.33 ± 6.38	-0.55 ± 2.94	-2.30 ± 3.82
in patients	6.82 ± 4.75	6.00 ± 6.32	5.98 ± 5.43	8.55 ± 6.78
p-values	< 0.001	< 0.026	< 0.001	< 0.001

Table 1. Results of Partitioning (Mean $\pm$ SD) using Bézier curves.

tested for difference between the two time points for all subjects using both methods as a way of comparing the consistency of calculating the PC ROI.

We then applied the Bézier curve method to 22 genetically or histologically proven AD patients and the 11 healthy (control) subjects. Scan interval was 491 ± 270 days. Cingulate masks were cleaved into thirds using Bézier curves, approximating each as the RAC, CAC and PC. Atrophy rates (mean % loss/year) of the whole cingulate and the three divided regions were calculated for all subjects.

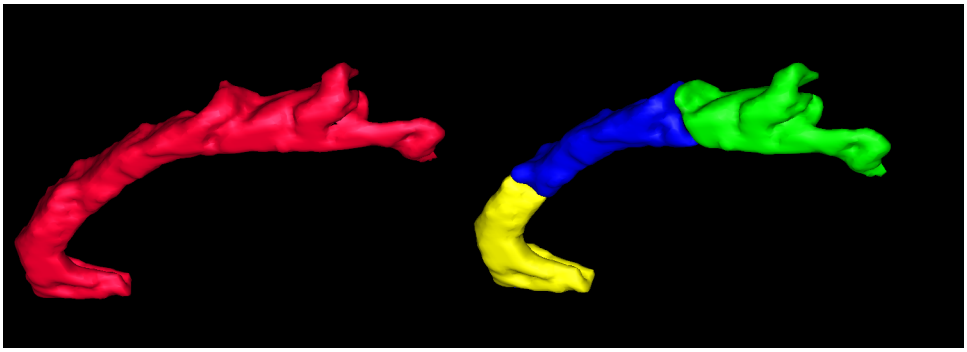


Figure 2. A three-dimensional view of a cingulate mask (left) and its three partitions (right). Orientation is $\uparrow$ superior, $\leftarrow$ anterior. The three partitions left-right are: RAC (yellow), CAC(blue) and PC(green).

3 Results

In our initial test of comparing the Bézier curve cleaving method with manual segmentation of the PC, we found that the difference in the variances between volume proportions at the two time points were signficantly reduced ($p < 0.01$).

The entire cingulate and all its sub-regions were able to differentiate between AD and control subjects (Table 1. Within the AD group the PC had a higher rate of atrophy than the cingulate, although it was not significant $(p > 0.1)$. This was due to the variance in the measurements which meant the cingulate was unable to differentiate AD subjects from controls any better than the whole cingulate. Also the PC had a higher rate of atrophy than that of other subdivisions, although this was not significant $(p > 0.06)$.

4 Discussion

The implication of a reduced $p-$value in the consistency test is that the Bézier curve cleaving method is more consistent than manual sub-division of the different sub-structures. This was because the method is fully automated and does not require any user intervention apart from the initial selection of start and end points and the selection of the fractional length. The Bézier curve method presented here is a consistent and reproducible way of partitioning a cingulate mask into different portions.

Within the AD group the PC had a higher rate of atrophy than the other subdivisions, in keeping with findings in the literature. Segmentation in this way gives us insight into the patterns of atrophy in the cingulate gyrus in AD. Improvements to the method's application could be made by performing the analyses on a more consistent scan interval and protocol, which may reduce the amount of variation in the measurements.

Fitting Bézier curves to the cingulate provides an accurate approximation to its central axis as the structure is long, reasonably thin and has a definite start and end point. Irregularities such as concavities and/or convexities are smoothed out as the ratio of these to the overall length of the structure are relatively insignificant. Any large

deformations in the structure would have an adverse affect on the final fit, however. It may be unsuitable for other structures if the central axis is not as apparent as that of the cingulate. An ideal structure for subdivision using this method would be the corpus callosum as it has a well defined central axis, and is also smooth.

References

1. D. J. A. Callen, S. E. Black, F. Gao et al. "Beyond the hippocampus." *Neurology* **57**, pp. 1669–1674, 2001.
2. H. Hampel, S. J. Teipel, G. E. Alexander et al. "Corpus callosum atrophy is a possible indicator of region and cell type specific neuronal degeneration in alzheimer disease." *Arch Neurol.* **55**, pp. 193–198, 1998.
3. P. A. Freeborough, N. C. Fox & R. I. Kitney. "Interactive algorithms for the segmentation and quantitation of 3d mri brain scans." *Comput. Method Programs Biomed.* **53**, pp. 15–25, 1997.

Registering 3D Lung Surfaces Using the Shape Context Approach*

Martin Urschler and Horst Bischof†

Institute for Computer Graphics and Vision, Graz University of Technology

Abstract. Studying the complex thorax breathing motion is an important research topic for medical (e.g. fusion of function and anatomy, radiotherapy planning) and engineering (reduction of motion artifacts) questions. In this paper we present first results on studying the 4D motion of segmented lung surfaces from CT scans at several different breathing states. For this registration task we extend the shape context approach for shape matching by Belongie et al. [1] from 2D shapes to 3D surfaces and apply it to segmented lung surfaces. Resulting point correspondences are used for a non-rigid thin-plate-spline registration. We describe our experiments on synthetic and real thorax data and show our quantitative and qualitative results.

1 Introduction

According to the European Respiratory Society, lung diseases rank second behind cardiac diseases in terms of mortality and cost of treatment. Computerized methods for objective, accurate and reproducible analysis of lung structure and function can provide important insights into these problems. However, due to the complexity of the breathing motion, investigations are often very complicated. In this paper we present first results on studying the 4D motion of segmented lung surfaces from several different breathing states scanned between Functional Residual Capacity (FRC) and Total Lung Capacity (TLC). We especially regard the problem of matching surfaces from consecutive breathing states and non-rigidly registering them by using a thin-plate-spline transformation model [2] for the deformation of the corresponding points. In general it is not possible to robustly derive corresponding features from the lung surfaces since the diaphragm-induced motion component and the movement of the rib cage tend to deform the elastic lung tissue, such that e.g. ridges might become valleys after deformation. The shape context approach introduced by Belongie et al. [1] was reported as a reasonable and promising method for matching 2D shapes (especially hand-written digits and letters) and 2D object recognition without relying on extracted features. We extend this approach to match 3D shapes and we are up to our knowledge the first ones to apply it to 4D medical image data, i.e. the segmented lung surfaces at several breathing states. Our image data stems from high-speed multi-detector spiral CT sheep studies. The sheep CT data was provided by Prof. Eric Hoffman, University of Iowa. The data is acquired at several (two, four or five) breathing states between TLC and FRC by a protocol where breath is held at fixed inspiration levels during the 30 sec scan time. This leads to a static breathing scheme, which has to be considered for the interpretation of derived motion models from matched and registered shapes. However, a protocol to scan thorax anatomy at different breathing states with high spatial resolution during dynamic (normal) breathing is currently not feasible. The image dimensions per breathing state are 512x512x550 with voxel dimensions of 0.52mm x 0.52mm x 0.6mm.

2 Method

2.1 Related Work

An older survey on the state of the art in 2D shape matching can be found in Veltkamp et al. [3]. Audette et al. give an algorithmic overview of surface registration techniques for medical imaging in [4], while Zitova et al. recently published an overview of image registration techniques [5]. Some examples for closely related methods for shape matching/registration are the modal matching approach proposed by Sclaroff et al. [6] or the TPS-RPM (Thin-Plate-Spline - Robust Point Matching) method developed by Chui et al. [7]. The main contribution of the work from Belongie et al. [1] is to present a robust and simple algorithm for finding shape correspondences by using shape context as a very discriminative representation that incorporates global shape information into a local descriptor.

2.2 The Shape Context Approach

The shape context approach [1] treats objects as (possibly infinite) point sets and assumes that the shape of an object is captured by a finite subset of its points, giving us a set $P = \{p_1, ..., p_n\}$. The points can be obtained as

* We gratefully acknowledge the support of Prof. Eric Hoffman, Department of Physiologic Imaging, University of Iowa, Iowa City for providing the CT image data.

†EMail: {urschler,bischof}@icg.tu-graz.ac.at

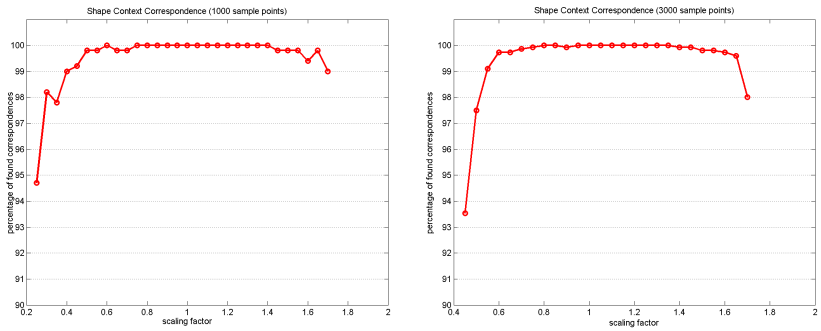


Figure 1. Percentage of found correspondences with the shape context approach. A series of scaling transformations was applied to a lung surface. Left plot shows results with 1000 sampled points, right image shows results with 3000 sampled points.

locations of edges from an edge detector or from another method to sample contour/surface points from a shape. The points need not and typically will not correspond to key points or structures such as maxima of curvature, inflection points or surface ridges. In contrast to the original implementation we lay strong emphasis on the discretization method. We are using a marching-cubes polygonization and sample contour points regularly from the constructed mesh. For each point p_i on the first shape, the "best" matching point q_i on the second shape has to be located. Therefore, the shape context descriptor is introduced. If we look at the set of vectors emitted from one point to all others, we can interpret this set as a rich description of the shape configuration relative to that point. Since this description is much too detailed, we take the distribution of the set of vectors as a compact, yet highly discriminative descriptor instead. So for each point p_i a histogram h_i of the relative position of the remaining points is calculated which is called the *shape context*. Now for point p_i from the first shape and q_j from the second shape, let $C_{ij} = C(p_i, q_j) = \frac{1}{2}\sum_{k=1}^{K}\frac{[h_i(k)-h_j(k)]^2}{h_i(k)+h_j(k)}$ denote the cost of matching these two points. Given the set of costs C_{ij} between all pairs of points on the first and second shape, we want to minimize the total cost of this one-to-one matching problem, which is an instance of the weighted bipartite matching problem. It can be solved in $O(N*(M+N*logN))$ time, with N being the number of nodes and M the number of edges in the graph. Here the original matching algorithm has been replaced by a more efficient one since the 3D case requires many more sample points than the 2D case which may lead to high run-times of the algorithm. The result of this step is a one-to-one mapping of corresponding points from the two shapes.

2.3 Non-Rigid Registration

After establishing the point correspondences we make use of the thin-plate-spline framework [2] due to its reported well suited applicability for modeling changes in biological forms. The thin-plate-spline approach leads to a transformation that consists of an affine part and a non-linear deformation part. The parameters of the thin-plate-spline model are calculated from the constraint that corresponding points are exactly interpolated and that the spline model between corresponding points is regular and smooth.

3 Results

To assess the validity of the shape context approach we performed quantitative evaluations on synthetic data sets. Further we used real thorax data for qualitative and quantitative evaluations. For our first synthetic study we took a single segmented lung surface and applied a series of known rigid scaling transformations to it. Afterwards we applied the shape context approach on the original and each transformed data set and calculated the percentage of correctly found corresponding points. The results for two distinct cases (sampling 1000 and 3000 points from original and transformed lung surface respectively) are depicted in Fig. 1.

Our second synthetic study uses a series of known thin-plate-spline transformations to create pairs of lung surfaces for shape matching. Therefore, we took two data sets at distinct breathing states (TLC and FRC) from our segmented real thorax data. We applied the shape matching algorithm and performed a thin-plate-spline transformation to get a plausible transformation T. This transformation T was then used as known "gold-standard"

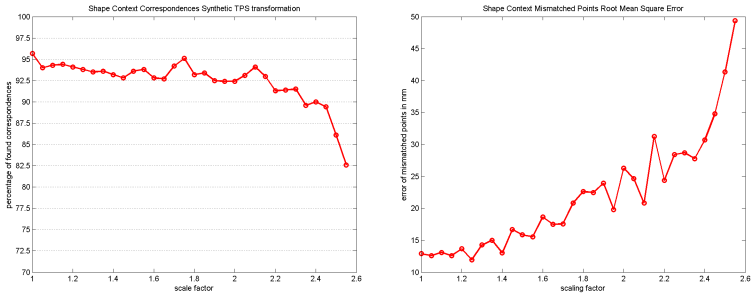


Figure 2. Matching results after applying a series of thin-plate-spline transformations with differing scales to a lung surface. Left plot shows percentage of found correspondences with the shape context approach on original and synthetically transformed data sets. Right plot shows the root mean square errors of the mismatched points compared to their original location. Surfaces were sampled with 1000 points.

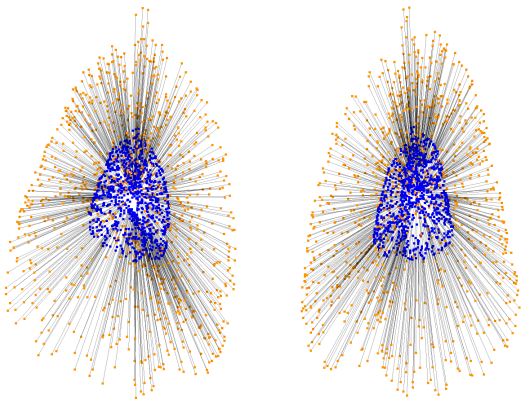


Figure 3. Two views of the matched points from the synthetically transformed (thin-plate-spline transformation) lung surfaces. The affine part of the thin-plate-spline transformation was a scaling transformation with scale factor 1.75.

transformation for our synthetic evaluation. We kept the non-linear part of T which resembles the deformations of the lung surface. Yet, the affine part of T was replaced by a series of scaling transformations with differing scale factors, giving a set of synthetic transformations $\{T_1, ..., T_n\}$. Then each of these transformations T_i was applied to the points of the lung surface at TLC. The TLC lung surface and the transformed lung surface were taken as input for the shape matching method and the resulting point correspondences were registered leading to transformations $\{T_1^{'}, ..., T_n^{'}\}$. So we could calculate a percentage of found correspondences and the root mean square error of distances between mismatched points and their original locations from comparing $T_i^{'}$ with T_i respectively. These results are depicted in Fig. 2. From the results of the second experiment we conclude that the performance of the shape context matching is very well in a wide range of scales, but decreases as soon as the scale factor gets too large. However, this is neglectable since too large scales resemble no meaningful simulation of breathing anymore. Two views of the matching result using a scale factor of 1.75 are shown in Fig. 3.

Finally we performed an evaluation using a data set with five different breathing states (TLC, FRC and three states inbetween). We built four subsets of these breathing states consisting of states {1,2,3}, states {2,3,4}, states {3,4,5} and states {1,3,5} respectively. The number of sample points was 1000 in all experiments. For each of these subsets we calculated the transformation $T^{'}$ relating first and second state and $T^{''}$ relating second and third states by using the shape context matching approach. Further we calculated $T^{'''}$ relating first and third state. Then we compared the results of applying $T^{'''}$ and applying $T^{''}(T^{'})$ on the first breathing state by calculating error statistics on the displacement of the transformed points. These results are given in Table 1. Screenshots of the results for a single pair of breathing states are shown in Fig. 4.

Data Subset	mean[mm]	std-dev[mm]	rms[mm]	min[mm]	max[mm]
{1,2,3}	5.81	4.61	7.42	0.14	24.24
{2,3,4}	5.97	4.39	7.41	0.29	25.57
{3,4,5}	6.18	3.87	7.29	0.21	22.10
{1,3,5}	6.14	4.17	7.42	0.12	35.02

Table 1. Results of the shape context validation on real data. For each subset mean, standard deviation, root mean square, minimim and maximum displacement of the transformed points was calculated.

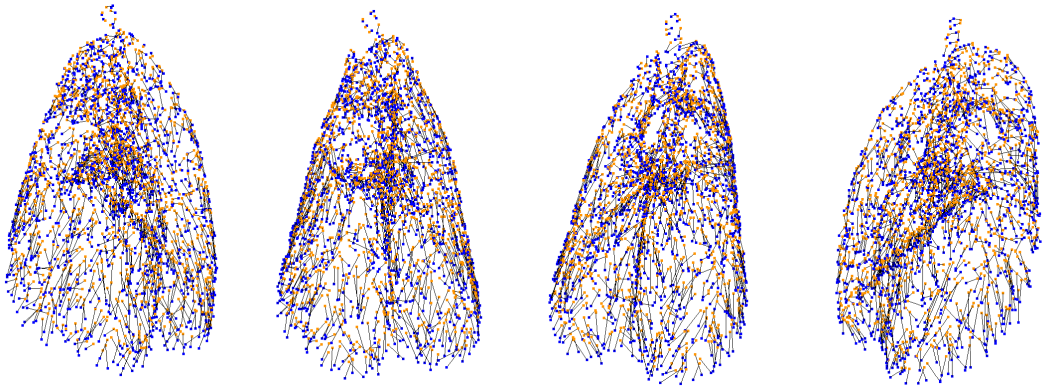


Figure 4. Four views of the matched points from an example pair of lung surfaces. 1000 sample points were used respectively for the shape context approach.

4 Discussion

We have demonstrated a 3D extension of the shape context approach for matching 3D lung surfaces. Shape context is a promising technique to find corresponding points for non-rigidly registering deformable anatomical structures which might also be useful for other 3D registration tasks in the medical domain. Our first experiments proved the validity on synthetic data and evaluated real-life data quantitatively and qualitatively. Future work will consist of more elaborate evaluations of the registration accuracy of our approach. In this context it will become possible to assess the semantic correctness of the found correspondences as well. Another intended task is a comparison of our approach with different state of the art matching techniques like the robust point matching approach proposed by Chui and Rangarajan [7]. Further topics will be to look into the elastic-body spline transformation proposed by Davis et al. [8] for registration instead of using the thin-plate-spline transformation and to investigate the robustness of this approach in case of noise and outliers.

References

1. S. Belongie, J. Malik & J. Puzicha. "Shape matching and object recognition using shape contexts." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24(4)**, pp. 509–522, 2002.
2. F. L. Bookstein. "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11(6)**, pp. 567–585, June 1989.
3. R. C. Veltkamp & M. Hagedoorn. "State of the Art in Shape Matching." Technical Report UU-CS-1999-27, Utrecht, 1999.
4. M. A. Audette, F. P. Ferrie & T. M. Peters. "An algorithmic overview of surface registration techniques for medical imaging." *Medical Image Analysis* **4**, pp. 201–217, 2000.
5. B. Zitova & J. Flusser. "Image registration methods: A survey." *Image and Vision Computing* **21(11)**, pp. 977–1000, October 2003.
6. S. Sclaroff & A. Pentland. "Modal Matching for Correspondence and Recognition." *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17(6)**, pp. 545–561, June 1995.
7. H. Chui & A. Rangarajan. "A new point matching algorithm for non-rigid registration." *Journal of Computer Vision and Image Understanding* **89**, pp. 114–141, 2003.
8. M. H. Davis, A. Khotanzad, D. P. Flamig et al. "A Physics-Based Coordinate Transform for 3-D Image Matching." *IEEE Transactions on Medical Imaging* **16(3)**, pp. 317–328, June 1997.

Eccentric Elliptical Contours in Total Hip Replacements

Stuart Kerrigan [a], Stephen J. McKenna [a], Ian W. Ricketts [a] and Carlos Wigderowitz [b] *

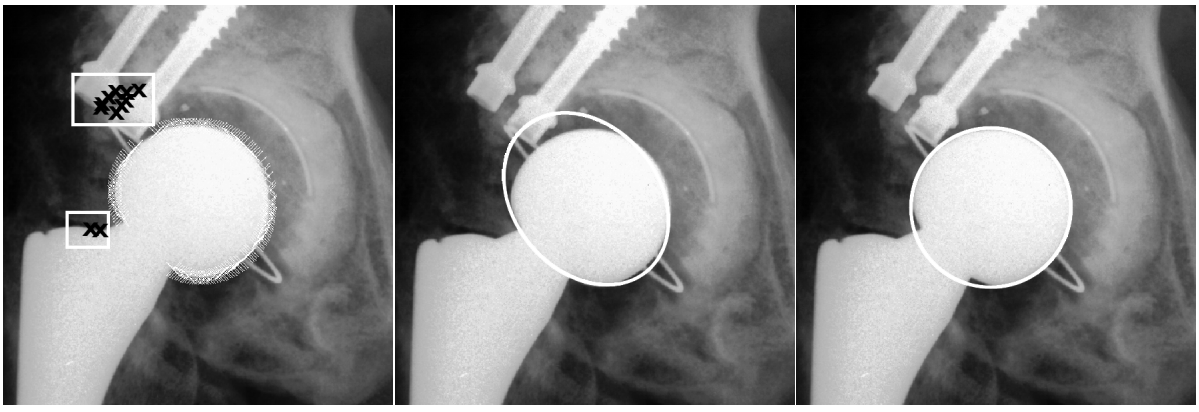
[a]Applied Computing, University of Dundee, DD1 4HN
[b]Orthopaedics and Trauma Surgery, Ninewells Hospital, Dundee, DD1 9SY

Abstract. The active ellipses method for assessing wear in total hip replacements uses robust ellipse fitting to localise the contours of the femoral head and acetabular rim wire marker. In the case of the latter these ellipses can be very eccentric and the standard algebraic distance was shown to be inadequate. The geometric distance from an ellipse to a point is not trivial to compute and thus numerous error of fit functions have been created. In this work several of these error of fit functions are compared, including a geometric error of fit function, on both synthetic data and by using active ellipses on a set of test radiographs containing eccentric rims.
Least squares estimation using a geometric error function was most accurate in the presence of Gaussian noise. However, least median of squares estimation using a geometric error function was most accurate in the presence of outliers. Furthermore, its performance was similar to that of a computationally cheaper error function known as the foci bisector distance to the extent that the two were almost interchangeable.

1 Introduction

The active ellipses method for assessing wear in total hip replacements (THRs) uses robust ellipse fitting to localise the femoral head and acetabular rim in radiographic images [1]. Radiopaque clutter, such as seen in Figure 1(a) causes structured outlying points from which a standard least squares (LS) ellipse fit [2] generates erroneous results, as seen in Figure 1(b). A robust method such as least median of squares (LMedS) is desirable in this instance as it has a breakdown point of 50% outlying points (see Figure 1(c)). Most conventional ellipse fitting uses an algebraic error function but this can cause problems when fitting eccentric ellipses such as those of the acetabular rim. However, other error of fit functions can be used, the most obvious choice being an error based on the geometric distance of a data point from the closest point on the ellipse curve. This distance is not trivial to compute and thus numerous computationally cheaper error functions have been considered in the past [3, 4]. Today's computational power and the availability of an efficient algorithm [5] have increased the feasibility of using the geometric distance.

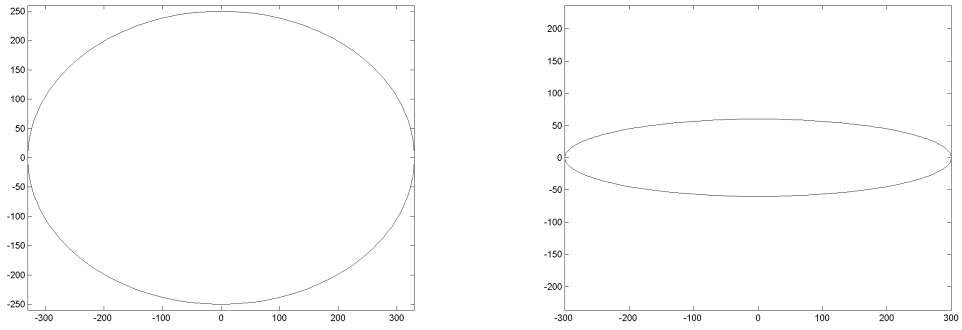
This paper reports an empirical comparison of the performance of LS and LMedS fitting with algebraic and geometric error functions using synthetic data. Additionally, the performance of LMedS with weighted algebraic and foci bisector distance error functions was evaluated. Finally, the best performing fitting algorithms were used to localise the elliptical structures in THR radiographs.



(a) (b) (c)

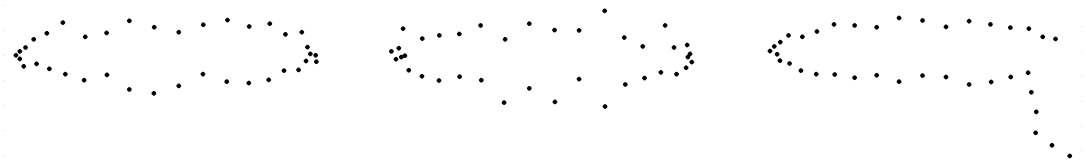
Figure 1. (a) Data points found during a femoral head search using an active ellipse. White crosses denote inliers. Outliers are shown in black and are highlighted by white rectangles. (b) An LS fit to the data points. (c) A robust LMedS fit which finds a good solution in the presence of outliers.

*{skerrigan,stephen,ricketts}@computing.dundee.ac.uk, cawigderowitz@lineone.net



(a) (b)

Figure 2. The two ellipses used to generate synthetic data. Both were centred at the origin and aligned with the image axes. (a) The less eccentric ellipse, $a = 333$, $b = 250$. (b) The more eccentric ellipse, $a = 300$, $b = 60$.



(a) (b) (c)

Figure 3. Data points created from the more eccentric ellipse. (a) Gaussian noise ($\sigma = 10$). (b) Half low variance Gaussian noise ($\sigma = 5$) and half high variance Gaussian noise ($\sigma = 20$ in this visualisation). (c) Structured outliers sampled from noisy line segments.

2 Materials and methods

2.1 Experiments using synthetic data

In order to assess the performance of the ellipse fitting algorithms, synthetic, noisy data sets were created from known ellipse parameters. Two ellipses were considered with eccentricities of 0.66 and 0.98 (see Figure 2). Eccentricity is defined as $\sqrt{1 - \frac{b^2}{a^2}}$ where a is the major semi-axis and b is the minor semi-axis of the ellipse. Three types of data set were created from each of the two ellipses.

Gaussian noise Data points were sampled at uniform intervals along the ellipse with additive Gaussian noise in the direction normal to the ellipse contour. Data sets were created with $\sigma = 0$, 10, 20, 30, 40, 50, 60 and 70.

Gaussian outliers Some points were sampled with noise drawn from a low variance Gaussian ($\sigma = 5$) while the remainder had high variance Gaussian noise ($\sigma = 500$) thus creating outlying points. Data sets were created with the percentage of points with high variance noise set to 0%, 10%, ..., 90%.

Structured outliers This type of data set was designed to simulate structured noise by sampling some of the points from straight line segments close to the ellipse. All points were sampled with Gaussian noise ($\sigma = 5$) from a closed contour, 80% from an elliptical arc, 10% from a line segment orthogonal to that arc, and 10% from another line segment rotated $45°$ with respect to the first line segment. Data sets were created with the percentage of structured outliers set to 0%, 10%, ..., 90%.

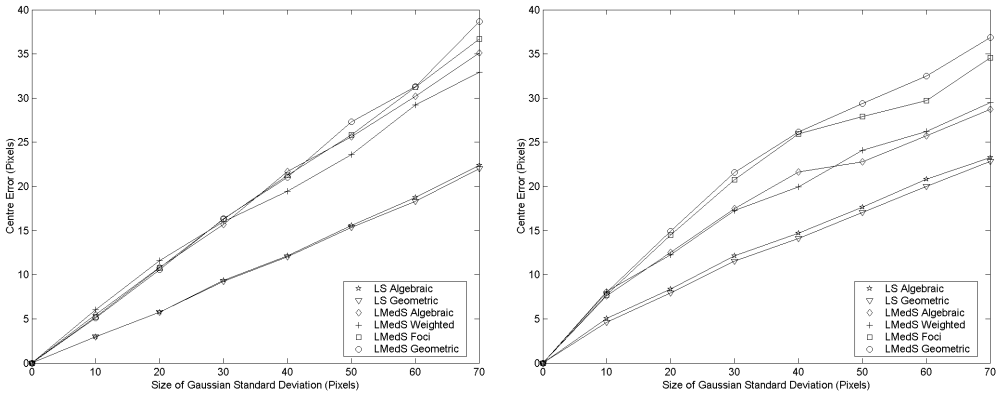
An example from each type of data set is shown in Figure 3. Each example consisted of 38 points. LS fits using geometric and algebraic error functions and LMedS fits (with LS fine tuning on resulting inliers) using algebraic, weighted algebraic by gradient [4], foci bisector distance [4] and geometric error [5] functions were performed. The Euclidean distances between the original and recovered centre points were used as a measure of accuracy.

2.2 Experiments using radiographic data

The elliptical projections of acetabular rims in standard clinical radiographs typically have eccentricities between 0.8 and 1.0. A set of 19 radiographs containing Zimmer CPT prostheses with particularly eccentric rim projections (> 0.96) was obtained. The most accurate error-of-fit functions from the experiments using the synthetic data sets described above were selected. These were used to perform active ellipse localisation on the radiographs using robust LMedS fitting. Localisation was run twice on each image, providing 38 results per method. No LS fine tuning was performed on the resulting inliers.

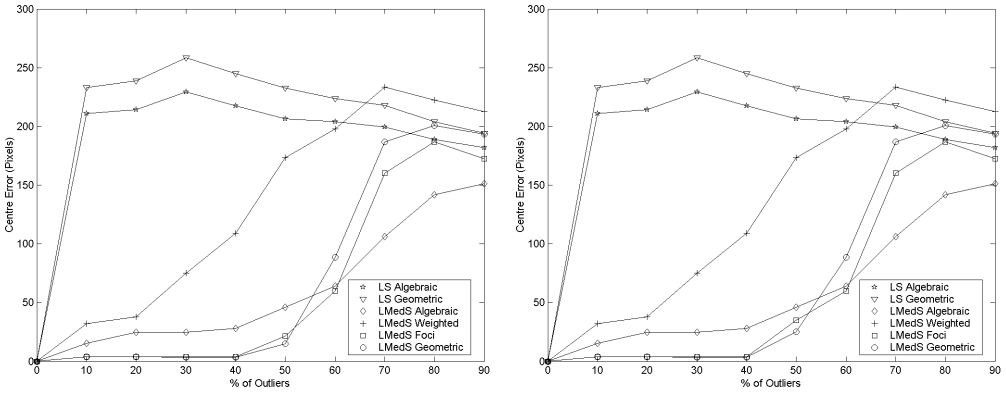
3 Results

Figures 4-6 show the alpha trimmed means ($\alpha = 0.1$) of the centre errors for each of the synthetic data sets. Each point on these plots was computed from 500 examples. In Figure 6 the centre error goes beyond the scale of the graph as selecting a large proportion of points from the clutter resulted in extremely eccentric and erroneous ellipses being generated.



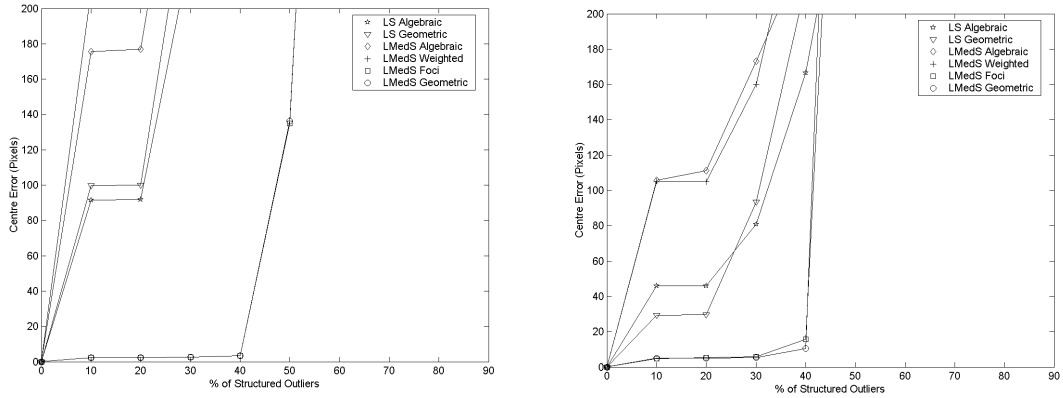
(a) (b)

Figure 4. Centre errors for (a) less eccentric and (b) more eccentric synthetic ellipse data as a function of σ.



(a) (b)

Figure 5. Centre errors for (a) less eccentric and (b) more eccentric synthetic ellipse data with Gaussian outliers.

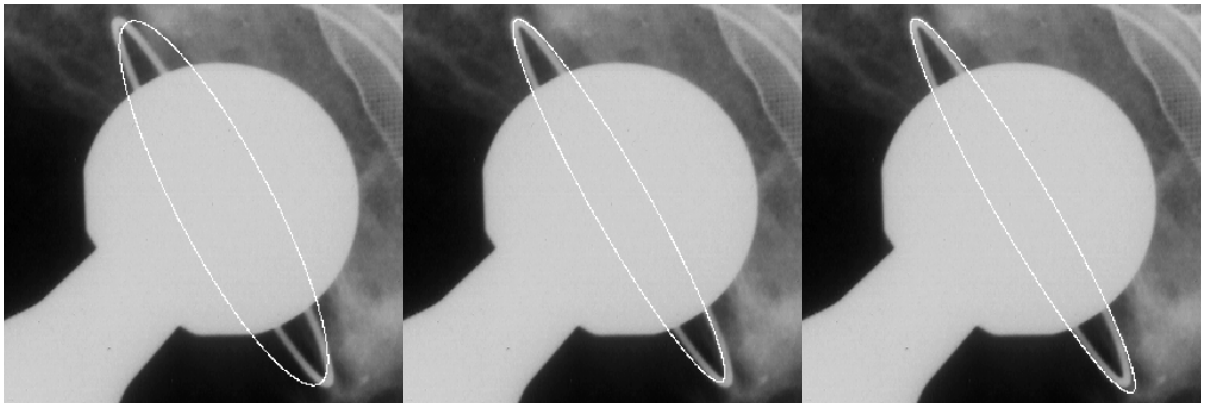


(a) (b)

Figure 6. Centre errors for (a) less eccentric and (b) more eccentric synthetic ellipse data with structured outliers.

On the radiograph dataset, geometric fitting failed 15 times out of 38, foci bisector fitting failed 16 times and algebraic fitting failed 32 times. Given the difficulty of the data set and the absence of any LS fine tuning to inliers identified from the minimal subset (which increases the performance of all three error functions), the results using

LMedS geometric fitting were encouraging. An example of the output of each of these algorithms is shown in Figure 7.



(a) (b) (c)

Figure 7. An eccentric rim with (a) failed algebraic, (b) successful foci bisector and (c) successful geometric fits.

4 Discussion

In the presence of pure Gaussian noise, LS outperformed LMedS irrespective of the error-of-fit function used (see Figure 4). LS using the geometric error performed the best. Least squares fitting is based on an assumption of Gaussian noise and is optimal under these circumstances. On the more eccentric ellipse the difference between LS algebraic and LS geometric became more pronounced, with LS geometric performing better. The LMedS geometric and LMedS foci bisector methods were least accurate in this case.

However, the robust LMedS fitting was more accurate in the case of outliers, whether Gaussian or structured. The plots in Figures 5 and 6 show that LMedS using foci bisector or geometric error-of-fit functions performed best on both ellipse eccentricities. The only noticeable difference between these two methods before the breakdown point was at 40% structured outliers in Figure 6(b). Results obtained with greater than 50% outliers lie beyond the theoretical break-down point of LMedS and so not surprisingly are poor. LMedS fitting with algebraic and or weighted algebraic functions performed very poorly in the presence of structured noise (see Figure 6). In fact LS methods were better in this case as LS fitted to both the straight line segments and the elliptical arc, while LMedS algebraic and weighted algebraic tended to favour points on the line segments.

The LMedS methods using geometric, foci bisector and algebraic functions were compared on the radiographic data. The latter was included because it has been used previously for this application. LMedS geometric made the most successful estimates, just outperforming LMedS foci bisector. LMedS algebraic performed poorly.

These experiments demonstrated that in the presence of non-Gaussian noise, LMedS geometric fitting tended to perform slightly better than the foci bisector distance approximation. However, it is a computationally expensive method and in applications where speed is important the foci bisector distance is recommended as an error of fit function. The authors are not aware of any previous work in the literature using the geometric distance [5] as an error function for LMedS ellipse fitting.

Acknowledgements: The authors acknowledge the support provided by Bristol Myers Squibb and Zimmer. They are grateful to Paul Rosin and Sung Joon Ahn for providing code and advice.

References

1. S. Kerrigan, S. J. McKenna, I. W. Ricketts et al. “Analysis of total hip replacements using active ellipses.” *Medical Image Understanding and Analysis* pp. 177–180, 2003.
2. A. Fitzgibbon, M. Pilu & R. Fisher. “Direct least square fitting of ellipses.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21(5)**, pp. 476–480, 1999.
3. P. Rosin. “Assessing error of fit functions for ellipses.” *Graphical models and image processing* **58(5)**, pp. 494–502, 1996.
4. P. Rosin. “Analyzing error of fit functions for ellipses.” *Pattern Recognition Letters* **17**, pp. 1461–1470, 1996.
5. S. J. Ahn, W. Rauh & H. J. Warnecke. “Least squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola and parabola.” *Pattern Recognition* **34**, pp. 2283–2303, 2001.

Feasibility study of a novel technique for constructing respiratory motion models, for use in 4D lung cancer radiotherapy planning

J R McClelland[a1], A G Chandler[a], J M Blackall[a], S Ahmad[b], D B Landau[b] and D J Hawkes[a]

[a] Division of Imaging Sciences, King's College London, Guy's Hospital, London UK,

[b] Department of Radiotherapy, Guy's and St. Thomas' NHS Trust, London UK

Abstract. Respiratory motion causes many problems in radiotherapy treatment planning for tumours in the lung. We have developed a novel method of constructing subject specific motion models that model and predict the complex 4D motion and deformation in a region of interest surrounding the tumour over the respiratory cycle. The model is built from several 'slabs' of CT data acquired rapidly in cine mode, during Free Breathing (FB), and a reference CT volume acquired at Breath Hold (BH). The reference volume is non-rigidly registered to each of the FB datasets. The model is built by interpolating over time between the non-rigid transformations that result from the registrations, and enables a prediction of what would be imaged in the region of interest at any arbitrary Position in the Respiratory Cycle (PRC). Presented here is a description of the method used to make the motion models, and an initial assessment of the feasibility of the methods by expert visual inspection of the results from two patients.

1 Introduction

Respiratory motion is a major factor contributing to errors and uncertainties in tumour localisation when planning radiotherapy treatment of lung cancer patients. As lung tumours can exhibit substantial movement over the respiratory cycle, and it has been shown that the shape of this movement does not vary significantly over time [1], it will be extremely beneficial to incorporate prior knowledge of this movement into radiotherapy planning. It will also be advantageous to account for the movement and deformation of the surrounding healthy lung tissue in order to minimise the irradiation of healthy tissue. Compensating for the respiratory motion can be achieved by gating the radiotherapy treatment with a respiratory signal, activating the beam only when the motion model used for planning indicates that the tumour is in the correct location. Alternatively the motion model and respiratory signal can be used to attempt to track and follow the tumour movements with a robotic radiosurgery system [2].

In previous work carried out by members of our group, motion models of the deformation of the liver over the respiratory cycle have been constructed [3]. These were constructed by interpolating over time between the non-rigid transformations that resulted from registering a reference MR volume to a series of MR volumes acquired throughout the respiratory cycle. The difference between this work and the method presented here, other than the choice of organ and clinical application, is that the series of MR volumes, although acquired at different PRC, were all acquired at BH, requiring the patient to suspend breathing at different PRC. However, it is both difficult for most people to suspend their breathing at a variety of PRC, and it has been shown that there are substantial differences between the shape and position of anatomy in BH and FB volumes [4]. In later work the constructing of motion models of the lung from rapidly acquired FB MR volumes was attempted [4]. Limitations in the quality of such volumes prevented the non-rigid registrations, and only affine registrations and modelling was possible.

There have been several recently proposed methods of creating 4D CT volume sets for use in radiotherapy planning [5,6]. Some of these use a cine acquisition mode [5] similar to that used in our method presented here, and some use a slow helical acquisition [6] and a modified reconstruction algorithm. However, they all differ from our method in that they attempt to produce a set of actual volumes, whereas once the model has been built our method only uses one (reference) volume, and deforms this using non-rigid transformations.

2 Method

2.1 Data acquisition

In order to construct our motion models it is necessary to obtain a high quality reference CT volume. This will be registered to the FB volumes to construct the model and deformed by the non-rigid transformations to produce the model's predictions. As the reference volume is not used in modelling the respiratory motion and

[1] Correspondence to: james.mcclelland@kcl.ac.uk

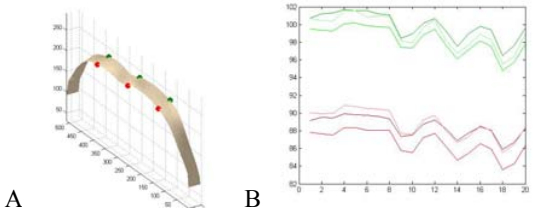
needs to cover the entire lung it is acquired at BH, at full exhalation, as this has been shown to be the most reproducible PRC. As well as the reference volume, smaller volumes acquired during FB using the CT scanner's cine mode are required. When volumes are acquired in the cine mode they are acquired and reconstructed with the couch in a fixed position. This enables rapid acquisition of a series of volumes, but the number of slices that can be simultaneously acquired (the size a 'slab' of cine data) is limited by the capabilities of the scanner. As the size of the slab of data is unlikely to cover the extent of the tumour movement and the region of interest, three contiguous slabs are acquired in adjacent couch positions. For each slab a series of 20 volumes are acquired over 15 seconds (0.5s x-ray time, 0.25s x-ray off time per volume), each volume consisting of 16 slices.

2.2 Non-rigid registration

The lung is segmented from the reference volume so that regions that are not expected to move with the lung are excluded. The registrations are performed using a modification of a non-rigid registration algorithm based on free form deformations using B-splines [7]. The reference volume is initially registered using an affine transformation to the FB volumes in order to align the reference volume with the different slab positions, and to account for some of the global motion due to respiration. The result of the affine registration is then used as an initial estimate for the non-rigid registration. The non-rigid registration is performed using a 17 x 17 x 7 Control Point Grid (CPG) with 21.83mm x 21.83mm x 3.75mm control point spacing.

2.3 Tracking the Position in the Respiratory Cycle (PRC)

It is necessary to assign each FB volume a PRC so that the results of the non-rigid registrations can be temporally interpolated over the respiratory cycle. To achieve this, the skin surface is automatically extracted from the FB volumes, by performing simple thresholding in the Anterior-Posterior (AP) direction. This gives a height for the skin (the AP co-ordinate) at every location specified by the Left-Right (LR) and Cranial-Caudal (CC) co-ordinates. Preliminary comparisons of the traces of the height of the skin, at 3 locations on the top and bottom slice of each slab show that there is no noticeable phase shift across each of the slabs, Figure 1. Hence the sum of the height of all locations of the extracted surface is used to increase the signal to noise ratio. Any regions of the traces that exhibit irregular breathing are excluded, and a sine wave is fitted to the remaining data points using non-linear least squares. The phase of the sine wave corresponding to each of the FB volumes is used to assign a PRC. As there is no observable phase shift across the slabs and the slabs are contiguous to each other, it is reasonable to assume there is no phase shift between the slabs, and therefore the PRC assigned to one slab will be implicitly in phase with the PRC assigned to another slab.



A B

Figure 1. A – Extracted skin surface showing the 6 locations that were tracked, B – the traces produced.

2.4 Constructing the motion model

The motion model produces a prediction of what would be imaged over the entire region of interest at any arbitrary PRC. This is constructed by combining the predictions of what would be imaged in each slab at the specified PRC. The reference volume is deformed with a non-rigid transformation to produce the prediction for each slab. We then simply translate the slab predictions into their correct locations with respect to each other. The non-rigid transformations used to produce the predicted 'slabs' at an arbitrary PRC are calculated by temporally interpolating between the original non-rigid registration results.

In order to temporally interpolate between the non-rigid registration results, each degree of freedom of the transformations (that is the x, y, and z displacement for each control point, so for the CPG used there are 17 x 17 x 7 x 3 = 6069 degrees of freedom) are interpolated separately. For each of the selected FB volumes from a slab the value of the degree of freedom is plotted against PRC. A third order polynomial is then fitted to these data points using linear least squares, which relates the value of the degree of freedom to the PRC. The coefficients of the polynomial are calculated for each of the degrees of freedom and can then be used to predict the value of

each of the degrees of freedom, and hence the entire transformation, that will deform the reference volume to predict what would be imaged in the relevant slab at any PRC.

3 Results and evaluation

The method of constructing motion models outlined above was attempted using data from two patients, one with a left lower lobe para-vertebral tumour and the other with a right upper lobe tumour. The non-rigid registration of the reference volume to the FB volumes was assessed by expert visual inspection, and for both patients was judged to be high quality. The boundary of the lung, the boundary of the tumour, and other visible anatomical structures in the lungs (blood vessels, lung fissures etc.) all showed good matches between the deformed reference volume and the original FB volumes.

Examination of the extracted skin surfaces demonstrated that the simple thresholding was very effective at extracting the skin surface, as there were no noticeable locations of any of the FB volumes where the skin had been incorrectly detected. Fitting the sine wave to the traces of skin height in order to assign a PRC worked with varying degrees of success. The traces from two of the slabs from patient 1 and one slab from patient 2 had regions that exhibited irregular breathing and needed to be excluded, and although the sine waves generally fitted the remaining data well, there are some points where the fit is not as good, Figure 2 (A: 10, 16, 18. B: 4, 16).

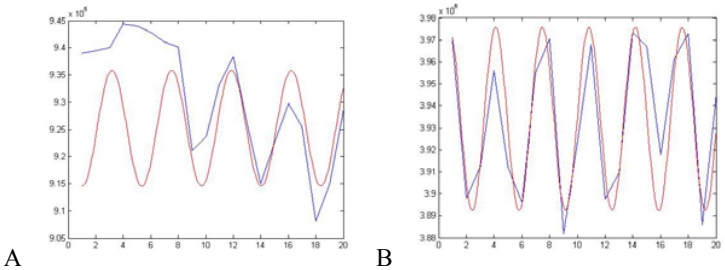


Figure 2. Traces of the extracted skin height, plotted against acquisition no. (time). Trace A shows how some regions exhibit irregular breathing and need to be excluded

The ability of the motion model to predict the slabs individually was assessed by producing a predicted volume, using the PRC assigned to each of the original FB volumes, and comparing the real and predicted volumes by expert visual assessment. Again, a number of anatomical features were used to assess the results, and although the lung boundary did not match in the two volumes as well as it did in the original registration results, the boundary of the tumour and other anatomical features demonstrated a very good match between the predicted and actual volumes, as shown in Figure 3.

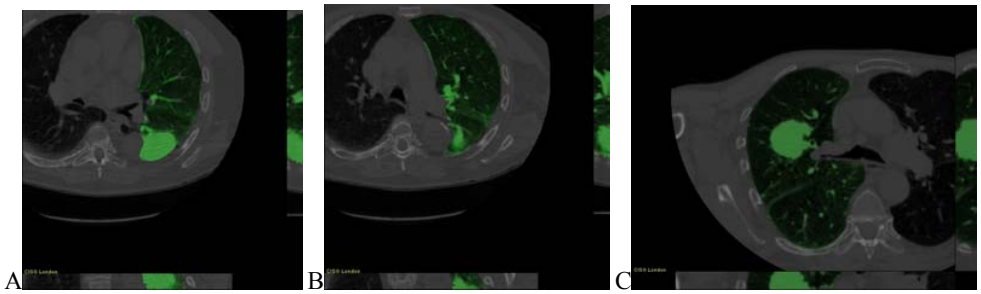


Figure 3. Examples of the deformed reference volume (green) overlaid on the original FB volumes.
A: patient 1 – slab 2 FB volume 7, B: patient 1 – slab 3 FB volume 9, C: patient 2 slab 2 FB volume 10.

It is more difficult to assess the ability of the motion model to predict the entire region of interest, as it is impossible to obtain good quality data over this whole region at the same time. Hence, we assessed the overall performance of the model by producing predictions at several equally spaced PRC and examined the discontinuities in the lung and tumour boundaries and other anatomical structures. Overall there were very few discontinuities across the three slabs for either patient at any PRC, and both patients displayed smooth continuous movement of the tumour and other features between slabs, as shown in Figure 4.

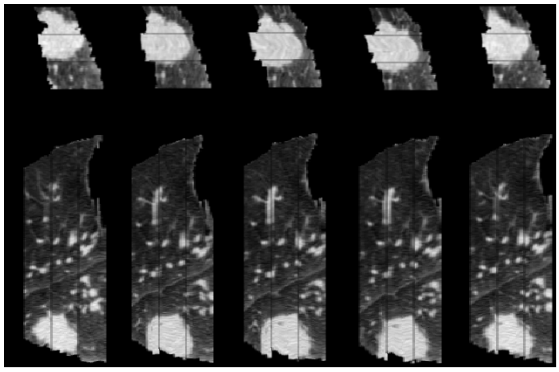


Figure 4. A series of predictions of the entire region of interest from 5 equally spaced PRC. Note the continuity of anatomical structures from one slab to the next (slab boundaries have been indicated by a dark line).

4. Discussion and Further Work

The initial assessment we have performed indicates that the motion modelling technique described here has considerable potential for improving radiotherapy planning. Now that we have demonstrated the feasibility of our method of constructing motion models, it is necessary to evaluate the models in much more detail. Our group is currently working on developing means of thoroughly assessing and validating the motion models predictions and performance, and producing quantitative as well as qualitative measures of their success.

Although the results show there is much promise for this method, they have also indicated there are several potential refinements to the technique, many of which we plan to incorporate into future work. Data from a larger region of interest may be beneficial, and so more slabs of cine data could be acquired. Fitting a sine wave to the trace from the skin surface, in order to assign the PRC, assumes a constant frequency and amplitude of the respiratory cycle, over theslab acquisition period. This is probably not true but, due to the low temporal resolution of the cine acquisition, it is very difficult to account for changes in amplitude or frequency of respiration. Using an external device to record a respiratory signal at a higher temporal resolution (tracked skin markers, spirometry) should facilitate a more accurate PRC being assigned. Functions other than polynomials need to be investigated for use in the temporal interpolation, as polynomials are not the ideal functions. Issues of reproducibility of respiratory motion and amount of data required to model it also need to be investigated.

Acknowledgements

The authors would like to thank Dr Shelia Rankin, Consultant Radiologist, and all of the CT Radiographers at Guy's Hospital for their valuable time and help. JRM receives funding from Vision RT and EPSRC via an Imaging Faraday Studentship, AGC receives funding from Philips, and JMB and SA receive funding from the Jill Calloumanopolous charitable fund.

References

1. Y. Seppenwoolde, H. Shirato, K. Kitamura, et al. "Precise and real-time measurement of 3D tumour motion in lung due to breathing and heartbeat, measured during radiotherapy", *Int. J. Radiation Oncology Biol. Phys.* **53**, pp 822-834, 2002
2. A. Schweikard, G. Glosser, M. Bodduluri, et al. "Robotic motion compensation for respiratory movement during radiosurgery", *Computer Aided Surgery* **5**, pp 263-277, 2000
3. J.M.Blackall, G.P. Penney, A.P. King, et al. "Tracking alignment of sparse ultrasound with preoperative images of the liver and an interventional plan using models of respiratory motion and deformation", in *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*. pp 218-227, Proceedings of SPIE Vol. 5367, Bellingham, 2004
4. J.M. Blackall, S. Ahmad, M. Miquel, et al. "Modelling respiratory motion for optimisation of lung cancer radiotherapy using fast MR imaging and intensity-based image registration", in *Proceedings of the International Society for Magnetic Resonance in Medicine*, **12**:2610, 2004
5. D.A. Low, M. Nystrom, E. Kalinin, et al. "A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing", *Medical Physics* **30**, pp 1254-1263, 2003
6. P.J. Keall, G. Starkschall, H. Shukla, et al. "Acquiring 4D thoracic CT scans using a multislice helical method", *Physics in Medicine and Biology* **49**, pp 2053-2067, 2003
7. D. Rueckert, L. Sonoda, C. Hayes, et al. "Nonrigid registration using free-form deformations: applications to breast MR images", *IEEE Transactions on Medical Imaging* **18**, pp 712-721, 1999

A comparison of serial MR neuroimaging at 1.5T and 3T

B Sneller[a], J Vickers[c], MA Dresner[b], RA Heckemann[b], J Fitzpatrick[b], ME Miquel[a], JV Hajnal[b], DLG Hill[a*]

[a]Imaging Sciences, KCL, [b]Imaging Sciences Department, Imperial College and [c]FMRIB, University of Oxford.

1 Introduction

Serial MR studies that seek to identify subtle changes over time typically involve registration and subtraction of the images from different time-points, followed by visual inspection or quantitative analysis of the difference images [1,2,3]. It is desirable to transfer serial MR imaging research protocols from one scanner to another, or to aggregate results from studies acquired on different scanners, for example for multicentre studies. With the increasing availability of MR scanners with a 3T field strength, it is important to study the consequences of this field strength difference on serial MR studies. We therefore compare serial magnetic resonance (MR) image registration results from similar acquisitions at 1.5T and 3T on 5 volunteer subjects.

2 Method

2.1 Image acquisition

Images were acquired on a 1.5T and a 3T Philips Intera scanner (Philips Medical Systems, Best, The Netherlands). The MR sequence was matched as closely as possible between scanners. A 3D gradient echo sequence (MPRAGE, flip angle of 8°, 1.2mm cubic voxels, readout AP) had the inversion time (TI) optimised, within the constraints of the acquisition, to give similar contrast (TI = 900 ms at 1.5T, TI = 1250 ms at 3T). Scans from 4 subjects were acquired using a birdcage coil at both field strengths (transmit receive at 3T, receive only at 1.5T). The remaining subject was scanned using an eight-element array coil at both field strengths. Using the array coils, a reference scan was used to correct for RF inhomogeneity (SENSE factor 1). Images were acquired, twice at each field strength, from five volunteers.

2.2 Image analysis

Each pair of images (a baseline and repeat scan acquired from a single subject at the same field strength) was rigidly registered by optimisation of normalised mutual information using the vtkCISG software package with 3 resolution levels, 64 bins and a maximum of 100 iterations. (www.image-registration.com). Difference images were produced by subtracting the repeat scans, which had been transformed using a Hanning windowed sinc interpolation (total width 12 voxels), from the baseline scans. The difference images were re-formatted in the coronal plane and the slice with the most severe flow artefact visually identified from each scan pair at each field strength. The artefacts in this slice were then assessed both quantitatively and qualitatively.

Quantitative analysis: Two regions of interest (ROI) were then identified in the selected slice. The artefact ROI was placed over an area in the difference image with the most severe flow artefact and the reference ROI placed over an area where no artefact was suspected, as shown in Figure 1. The standard deviation of the difference values in each ROI was calculated, and the ratios of these standard deviations were then compared.

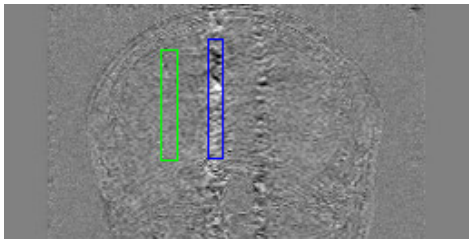


Figure 1. Measuring severity of flow artefact. Reference ROI (left) and artefact ROI (right)

Qualitative analysis: The artefacts in the reformatted coronal slice were visually assessed by an observer who was blinded to the field strength used. The difference images were all windowed to +/- 35 % of the cortical grey matter intensity (the mean intensity in similar regions of cortex in the baseline scan of each scan pair). The

*Author for correspondence: derek.hill@ieee.org

selected difference image slices acquired from the same volunteer with the same coil at 1.5T and 3T were paired together. The pairs of difference images were randomly ordered, and shown to the observer, alongside the corresponding coronal cross sections from the baseline scans. The observer was asked to view the pairs of difference images (one from 1.5T, the other from 3T, randomly ordered), and rank the artefacts on a 5-point scale (much worse, slightly worse, same, slightly better, much better).

3 Results

3.1 Images

Some of the acquired images and the generated difference images can be seen below.

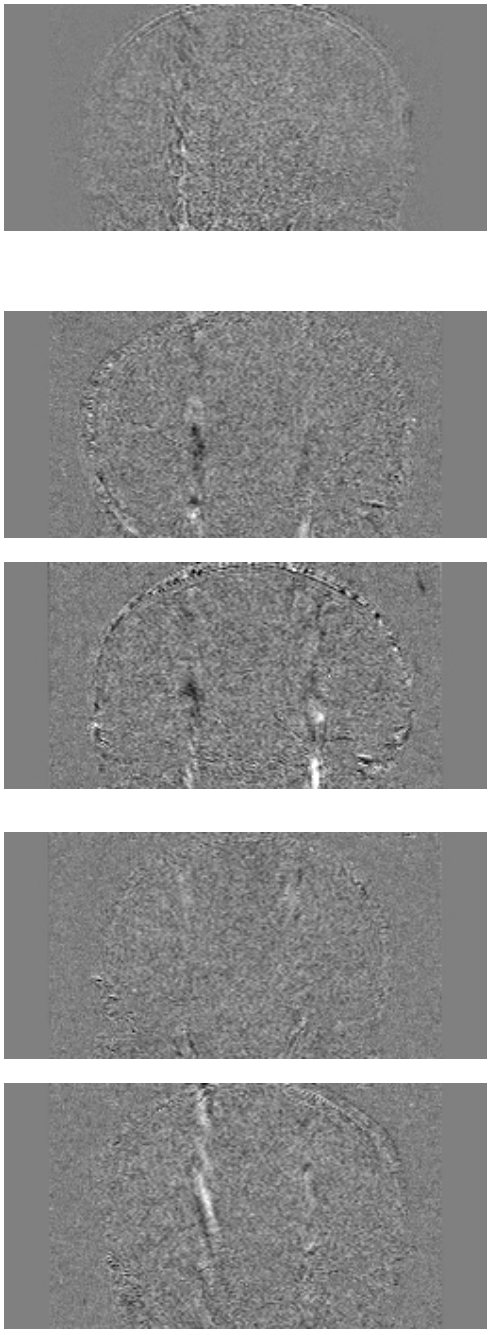


Figure 2. Subtraction images 1.5T

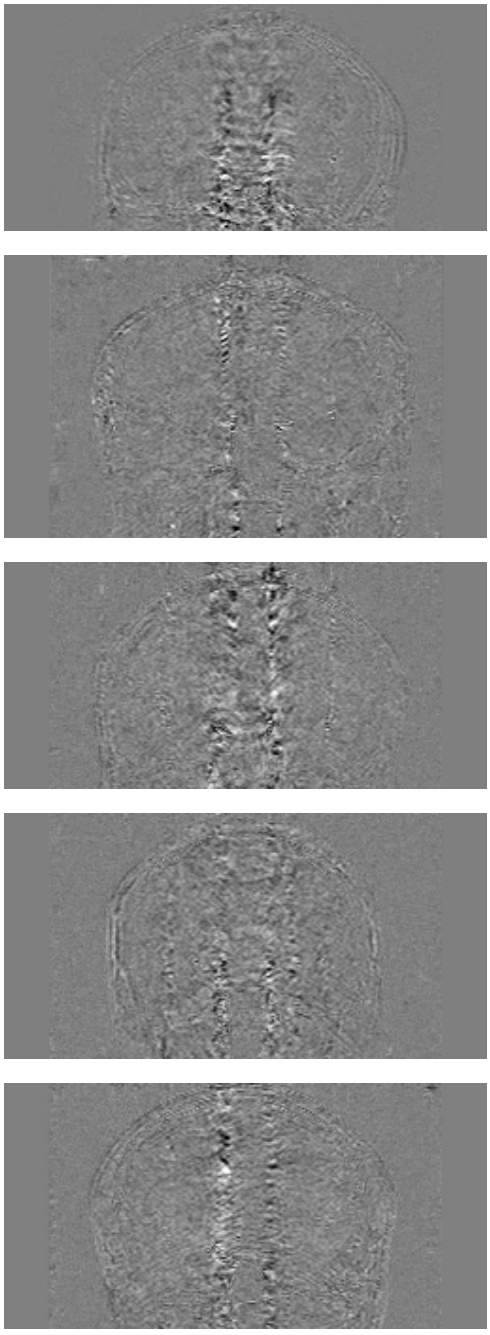


Figure 3. Subtraction images 3T

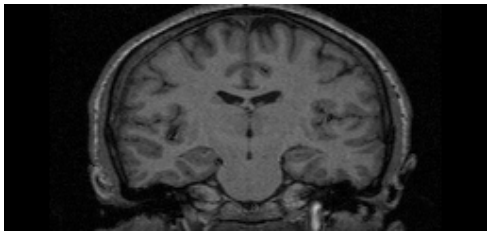


Figure 4. Coronal section 1.5T

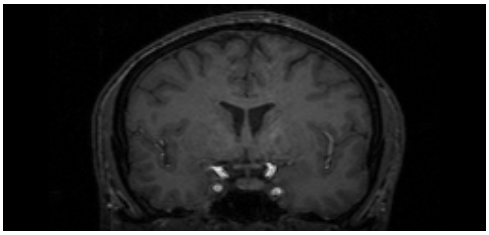


Figure 5. Coronal section 3T

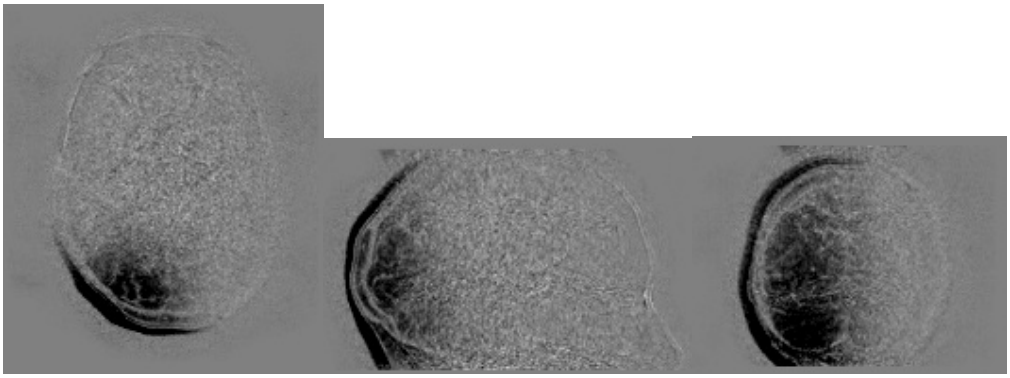


Figure 6. Example of array coil malfunction

3.2 Quantitative analysis

In the tables below are the ROI standard deviations measured from difference images produced from both the images acquired at 1.5T, see Table 1, and those from images acquired at 3T, see Table 2.

1.5T	Artefact standard deviation	Non-artefact standard deviation	Ratio of the standard deviations
Volunteer 1	18.79	9.79	1.919
Volunteer 2	17.00	11.37	1.495
Volunteer 3	18.50	11.31	1.636
Volunteer 4	13.27	10.60	1.252
Volunteer 5	21.57	11.60	1.859
Average	17.83	10.93	1.632

Table 1. Results from 1.5T difference images

3T	Artefact standard deviation	Non-artefact standard deviation	Ratio of the standard deviations
Volunteer 1	71.81	14.88	4.826
Volunteer 2	48.11	15.40	3.124
Volunteer 3	55.13	14.10	3.910
Volunteer 4	48.31	15.02	3.216
Volunteer 5	60.19	15.85	3.797
Average	56.71	15.05	3.775

Table 2. Results from 3T difference images

It is clear from tables 1 and 2 that the artefact level at 3T is larger than at 1.5T (more than twice as large on average), and this difference was found to be significant at the 5% level using a Student's t-Test ($p=0.0006$).

3.3 Qualitative results

An observer rated the images acquired from 3T as having the more pronounced flow artefacts in all 5 sets of images, with one set's flow artefacts rated much more pronounced at 3T. Also more of the coronal slices showed clear flow artefacts in the 3T difference images compared with the 1.5T difference images.

4 Discussion

Comparison of 1.5T and 3T images showed the expected differences in gradient calibration and contrast, see Figures 4,5. However, on inspection of the registered difference images, more pronounced flow artefacts could be seen in the 3T images. These artefacts were quantified and found to be significantly larger at 3T ($p<0.05$). At 1.5 T the resulting difference images are dominated by noise, see Figure 2, but at 3T they are dominated by flow artefacts that appear to arise from the carotid siphons, see Figure 3. In all of the 1.5T images more lateral artefacts dominate, while more medial artefacts were much less visible. However at 3T the most pronounced flow artefacts were more medial, although clear artefacts were also visible more laterally. These patterns were found using both birdcage and the array coils. For one array coil examination (not one of the five pairs referred to above), a single coil element malfunctioned at the time of the acquisition, and the difference image from this scan clearly shows the resulting artefact, see Figure 6.

5 Conclusion

This study has illustrated some of the difficulties of transferring image acquisition and analysis protocols from 1.5T to 3T, or carrying out studies that seek to aggregate data from 1.5T and 3T scanners. Further investigation is needed in order to quantify and minimize this effect.

Acknowledgements

This work is funded by the UK e-Science project IXI. We wish to thank Dr. Vivek Muthurangu for his assistance with this work.

References

1. Hajnal JV, Saeed N, Oatridge A, et al. "Detection of subtle brain changes using subvoxel registration and subtraction of serial MR images", *J Comput Assist Tomogr* **19**, pp 677-91,1995.
2. Freeborough PA & Fox NC "The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI" *IEEE Trans Med Imaging* **16**, pp 623-9,1997.
3. Denton ER, Holden M, Christ E, et al. "The identification of cerebral volume changes in treated growth hormone-deficient adults using serial 3D MR image processing", *J Comput Assist Tomogr* **24**, pp139-45, 2000.

An MR Neuroimaging Tissue Segmentation Study at 1.5T and 3T

J. Vickers[a], B. Sneller[c], J.V. Hajnal[b], D.L.G. Hill[c], R. Heckeman[b], S.M. Smith[a]

[a]FMRIB, University of Oxford [b]Imaging Sciences Centre, Imperial College [c]Imaging Sciences, KCL

1 Introduction

We investigate the performance of state-of-the-art segmentation tools when processing multi-site structural MRI brain data. T1 and dual-echo T2/PD data was acquired at both 1.5T and 3T with two coils (Birdcage & Array Coil) and varying SENSE factors. The effects of differing magnet field strengths and coil configurations on tissue classification was studied both qualitatively and quantitatively.

2 Method

Image acquisition All images were acquired on 1.5T and 3T Philips Intera scanners. Both machines had 8-element array coils and birdcage coils (transmit & receive at 3T, receive only at 1.5T). The MR sequences were matched as closely as possible between scanners. The 3D gradient echo sequence (MPRAGE, flip angle of 8 degrees, 1.2mm cubic voxels, readout AP) had the inversion time optimized, within the constraints of the acquisition, to give similar contrast. Using the array coils, a reference scan was used to correct for RF inhomogeneity (SENSE factor 1). Images were acquired from two volunteers, both imaged twice in each field strength and coil combination.

Image analysis Analysis was carried out using tools from FSL (FMRIB Software Library – www.fmrib.ox.ac.uk/fsl). The single and multi-channel segmentation tool used, FAST/mFAST [Zhang 2001], incorporates iterative bias field removal that is conducted alternately with the hard segmentation iterations. This bias field correction involves applying lowpass smoothing filters to the residual mean and mean residual covariance fields. The segmentation algorithm can make use of spatial segmentation priors. FAST and mFAST perform finite Markov Random Field mixture modelling to produce partial volume tissue-type estimates, which we used in the quantitative cerebro-spinal fluid, grey matter and white matter volumetric segmentation calculations. The workflow procedure is designed to minimize the variation due to the brain extraction (BET [Smith 2002], required as a pre-processing step before running FAST and mFAST) through the derivation of a single brain mask for each subject to be used to preprocess data from all scanners, coils, and acquisitions; optimal algorithm parameters are selected for a representative T1 image by an expert user. The brain mask for a given subject is then propagated to that subject's different scans through the use of a linear image registration tool, FLIRT [Jenkinson 2002].

Through such a workflow it is ensured that interpolation problems are removed as all segmentations are performed in native space. Given that different scanners have slightly different calibrations, BET is also used to extract the skull, which is used to give a normalisation (scaling) constant in the propagation of brain masks and comparisons of voxel counts.

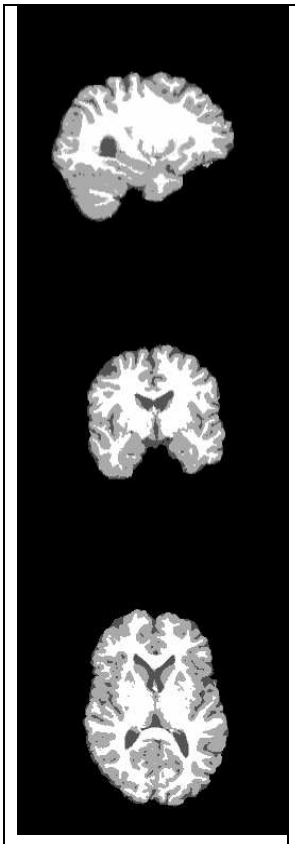
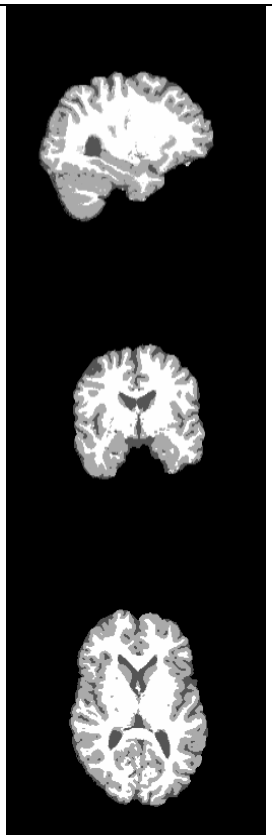
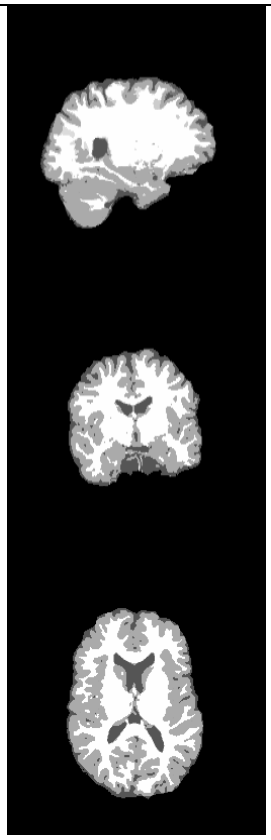
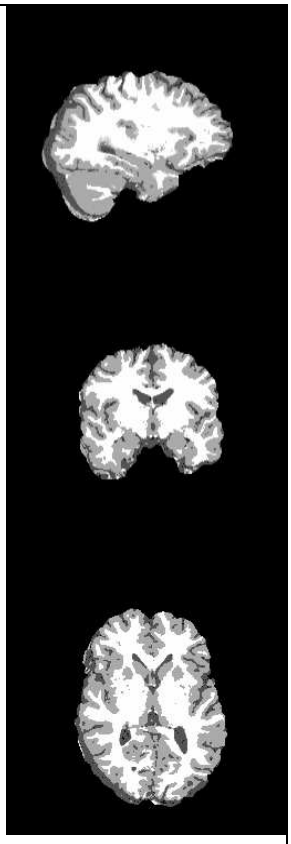
3 Results

In general the segmentations obtained using FAST from the T1-weighted gradient echo images for the same subject looked similar to each other visually but with some misclassification of true deep grey matter.

The segmentations of the T2/PD-weighted pair obtained using mFAST were quite dissimilar to the segmentations obtained from the T1-weighted images due to the differing physical spin relaxation processes being measured.

Individual voxel tissue classification differences could be observed visually.

Qualitative comparison of Segmented T1-weighted images for one subject

			
(A) 1.5T Birdcage	(B) Repeat of (A)	(C) 3T Birdcage	(D) 1.5T Array Coil

Quantitative Results

A Normal healthy subject has T1-weighted scans taken and repeated with Birdcage and Array Coils at both 1.5T and 3T

Subject A T1 Image	Scaling	Raw Csf mm3	**% Csf Scaled**	Raw Grey mm3	**% Grey Csf**	Raw White mm3	**% White Csf**	Total mm3	Scaled Total mm3
(A)Birdcage 1.5T	1.18403	204.1	**13.9**	723.9	**49.1**	545.3	**37.0**	1473.3	1744.4
(B)Birdcage 1.5T Repeat of (A)	1.18403	202.9	**13.8**	720.0	**48.9**	549.0	**37.3**	1471.9	1743.9
(C)Birdcage 3T	1.14533	240.0	**16.0**	719.2	**48.0**	539.9	**36.0**	1499.1	1716.9
(D)Birdcage 3T Repeat of (C)	1.14553	239.0	**16.0**	718.2	**47.9**	541.0	**36.1**	1498.2	1716.2
(E) Array Coil 1.5T	1.18403	256.0	**17.4**	662.9	**45.1**	552.4	**37.5**	1471.3	1742.1
(F)Array Coil 3T	1.14553	243.5	**16.2**	725.5	**48.2**	536.1	**35.6**	1505.1	1724.1

A second normal healthy subject has T1-weighted scans taken at both 1.5T and 3T with Birdcage coils and scan repeats.

Subject B T1 Image	Scaling	Raw Csf mm3	**% Csf Scaled**	Raw Grey mm3	**% Grey Csf**	Raw White mm3	**% White Csf**	Total mm3	Scaled Total mm3
(A)Birdcage Coil 1.5T	1.40029	179.8	**14.3**	632.2	**50.2**	448.2	**35.6**	1260.2	1764.6
(B)Repeat of (A)	1.40029	179.1	**14.2**	633.8	**50.3**	447.2	**35.5**	1260.1	1764.5
(C)Birdcage 3T	1.38896	200.0	**15.3**	663.7	**50.6**	446.8	**34.0**	1310.5	1820.2
(D)Repeat of (C)	1.38896	198.6	**15.2**	663.6	**50.6**	448.4	**34.2**	1310.6	1820.3

Discussion of Results

These preliminary results suggest that repeat scans on the same scanner using the same RF coil have closer agreement than scans acquired at different field strengths (1.5T vs 3T), or using different RF coils (birdcage vs. array). These differences between scanning technologies could confound inter-subject differences being studied. This could be for one or more of the following reasons: the images have different contrast to noise ratios, and different B1 inhomogeneity; the segmentation workflow procedure, which was designed carefully to reduce the amount of brain extraction necessary, may introduce a bias; the segmentation tool (FAST) used may introduce a bias. The relative contribution of these factors needs further investigation.

Major differences could be seen qualitatively in the segmentation images between T1-weighted and T2/PD-weighted dual-echo acquisitions of the same subject, particularly in the deep grey matter structures of the inner brain.

4 Conclusion

This study has illustrated some of the difficulties in transferring image acquisition and analysis protocols from 1.5T to 3T and between differing types of RF coil arrangements, or of carrying out studies that seek to aggregate data from such scanners. Further research is needed to minimize these issues in order to make it possible to carry out large-scale multicentre neuroimaging studies.

The quantitative study of segmentation results has been seen to involve wider image processing issues such as the confounding factors of the use of Brain Extraction tools (BET) as a pre-processing step, the required use of linear registration tools to carry brain masks to native space and its use in the scaling of voxel sizes using the normalisation factor of the skull. Further investigation is needed in order to establish the exact relationship in the changes observed in this study, and then to find methods to minimize or understand better such differences. We also intend to study the interaction of the various smoothing controls (i.e., smoothing extent of bias field and of segmentation labelling) in the segmentation process and the different kinds of bias field found at different field strengths and with different kinds of RF coils.

References

[Jenkinson 2002] Jenkinson, M., Bannister, P., Brady, J., and Smith, S. Improved optimisation for the robust and accurate linear registration and motion correction of brain images. NeuroImage, 17(2):825-841.
[Smith 2002] Smith, S. Fast robust automated brain extraction. Human Brain Mapping, 17(3):143-155.
[Zhang 2001] Zhang, Y., Brady, M., and Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm. IEEE Trans. on Medical Imaging, 20(1):45-57.

Acknowledgements This work is funded by the UK e-science project IXI.

Author for correspondence: john@fmrib.ox.ac.uk

Improved T1 estimation in contrast enhanced MRI for pathology detection using optimum flip angle acquisitions

Georgios Ketsetzis[a*] and Michael Brady[a]

[a]Medical Vision Laboratory, Department of Engineering Science, Robotics Research Group, Ewert House, Ewert Place, Summertown, Oxford OX2 7BZ

Abstract

In this paper we establish a new methodology for the selection of multiple optimum flip angle acquisitions prior to injection of contrast agent during breast screening, by extending our recently developed mathematical framework [1]. We formulate the problem using the standard MR signal model and perform error analysis assuming that errors in estimating T_1 are due to signal noise. Our objective minimization function depends explicitly on the noise model for the MR signal and the flip angle. We determine analytically the optimum flip angle which minimizes the error in estimation of T_1 values that are characteristic of tumor types. We use our analytic derivation of the optimum flip angle to provide an extension of the recent method by Armitage et al. [2] and Monte-Carlo simulations of real data to demonstrate that our protocol is significantly more accurate and more stable to noise with identical computational cost. The resulting T_1 maps obtained by the proposed method are proved to be more reliable and can therefore be used for improved diagnosis of breast disease.

1 Introduction

Dynamic contrast-enhanced magnetic resonance imaging (CE-MRI) is an MRI technique which assesses tissue properties. Contrast agent (typically Gd-DTPA) is injected into the patient immediately prior to acquiring a series of T_1 weighted MRI volumes using eg. Fast Spoiled Gradient (FSPGR) echo sequences, with a temporal resolution currently around a minute [3,4]. The presence of contrast agent within an imaging voxel results in an increased signal that can be observed during the time course of the experiment. Different tissue types have different contrast uptake properties and such signal-time curves enable their identification. Study of these curves has been used clinically to identify and characterize tumors into malignant or benign classes, although success has been variable with generally very good sensitivity ($> 95\%$) but often quite variable specificity.

The primary reason for the poor specificity is that pharmacokinetic modelling of uptake curves is based on the erroneous assumption that relative signal enhancement (RSE) is linearly proportional to contrast agent concentration. It has been shown [5] that this relationship is non-linear and that it can be represented as a function of two variables: T_{10}, the T_1 tissue relaxation time before injection of the contrast agent, and the concentration of the contrast agent. Given the RSE, an accurate measurement of T_{10} suffices for determination of the contrast agent concentration and more reliable classification of breast tissues according to the uptake curves of the pharmacokinetic model.

In a recent study on the measurement of T_{10}, Armitage et al. [2] developed a method based on Monte-Carlo simulations for minimizing the error in T_{10} estimation arising from signal noise with respect to the choice of flip angle. The method is feasible in determining up to 3 optimum flip angles for signal acquisition for the measurement of T_{10}, but cannot predict a fourth or subsequent flip angle acquisition in real time applications. In addition it has been demonstrated that an increasing number of flip angle acquisitions leads to increasingly accurate T_{10} values obtained after linear fitting of the data. In this paper we show that a determination of a fourth and subsequent flip angles is obtainable at no additional cost by extending the method of Armitage. This is achieved by a concise error analysis which leads naturally to the establishment of the relationship between the noise model for the signal and the set of flip angles which minimize the error in T_{10} estimation. Following the formulation of the problem, in Section 2, Section 3 provides the solution and a critical comparison of our proposed method against the method of Armitage et al. in terms of numerical stability, robustness and accuracy, using synthetic data. Monte-Carlo simulations demonstrate a clear advantage of our proposed method. This new method results in improved T_{10} mapping of breast tissue and can be used to enable more accurate diagnosis of breast cancer by enabling more reliable classification of breast tissue, especially in classifying tumour types.

2 Methodology for the measurement of longitudinal relaxation time and problem formulation

The standard model of the signal generated at a voxel by a gradient echo MR pulse sequence prior to injection of contrast agent, is given by:

$$S = g\rho e^{-\frac{TE}{T_{20}^*}} \sin(\alpha) \frac{1 - e^{-\frac{TR}{T_{10}}}}{1 - \cos(\alpha) e^{-\frac{TR}{T_{10}}}}, \tag{1}$$

where S is the measured signal, g is the scanner gain, ρ is the proton density, TE the echo time, TR the repetition time, α the flip angle and T_{10} and T_{20}^* are the longitudinal and transverse relaxation times respectively. From now on, we write

$$k = g\rho e^{-\frac{TE}{T_{20}^*}}. \tag{2}$$

*Corresponding Author's email: geok@robots.ox.ac.uk

Proposed methods for measuring T_{10} in current 3D breast imaging include inversion recovery protocols [6] and gradient echo sequences acquired with variable TR. Both are extremely time inefficient [2]. Instead, variable flip angle methods are the most viable option. Our focus in this paper is to examine which selection of variable flip angle acquisitions produce T_{10} measurements with the highest accuracy.

Multiple angle acquisitions can be used to determine T_{10} by fitting of the signal equation (1) to data. This is achieved by a re-arrangement of equation (1) such that:

$$\overbrace{\frac{S}{\sin\alpha}}^{Y} = \overbrace{e^{-\frac{TR}{T_{10}}}}^{A}\overbrace{\frac{S}{\tan\alpha}}^{X} + \overbrace{k(1-e^{-\frac{TR}{T_{10}}})}^{B}, \tag{3}$$

and T_{10} is obtained by the intersection of the line with the x-axis [7]. This method can be used to produce a resulting value for T_{10} for each voxel location which can then be represented as an image (T_{10} map). An increasing amount of signal acquisitions is a time consuming process that may prolong a scanning session and cause inconvenience to the patient, but also introduces corruption of the obtained scanner data by patient movement that must be compensated for during the data analysis. It is therefore crucial that an a-priori protocol for signal acquisitions is known so as to compensate between the increasing amount of signal acquisitions required for obtaining stable values for T_{10} and the time constraints imposed by the need for rapid screening with the least possible patient inconvenience. This protocol should consist of a reasonable amount of flip angle acquisitions which optimize the accuracy in the obtained value for T_{10} as predicted by fitting the obtained data to equation (3). The common problem encountered is the existence of noise in the signal acquisitions, due to magnetic field inhomogeneities, radio frequency pulse miss-calibrations or corruption of the MR signal by the resistance of the receiver coil of the scanner or the particular condition of the patient. Moreover, there is currently no criterion which can determine the optimum number of flip angle acquisitions, beyond which no further improvement in the value of T_{10} can occur. A proposed protocol for signal acquisition for accurate T_{10} mapping should express clearly the quantitative and qualitative relationship between signal noise and optimum flip angle selection, a task that has not been addressed so far.

3 Analysis of the error in the measurement of T_1

To study the error in the measurement of T_{10} using low flip angle acquisitions as described in Section 1, we suppose that at each voxel location $\vec{\mathbf{x}}$ there exists a random error to the signal due to noise denoted by $\mathcal{E}_S(\vec{\mathbf{x}})$ and a corresponding error for k denoted by $\mathcal{E}_k(\vec{\mathbf{x}})$. The following theorem provides the error in the measurement of the parameter T_{10} using a FSPGR echo pulse sequence. We are able to prove [1]

Theorem 3.1. *Let $\hat{S} = S + \mathcal{E}_S(\vec{\mathbf{x}})$ and $\hat{k} = k + \mathcal{E}_k(\vec{\mathbf{x}})$, where S, k are the real values of the signal and the parameter k respectively and $\mathcal{E}_*$ denotes the corresponding random errors. If T_{10} represents the actual value of T_{10} at $\vec{\mathbf{x}}$, the resulting value of T_{10} using a FSPGR echo pulse sequence is*

$$\hat{T}_{10} = T_{10} - \frac{\mathcal{W}_n}{\mathcal{W}_d} + \mathcal{O}(\max(\mathcal{E}_S, \mathcal{E}_k)^2),$$

where

$$\mathcal{W}_n = (\mathcal{E}_S(e^{\frac{TR}{T_{10}}} - \cos(\alpha)) - \mathcal{E}_k \sin(\alpha)(e^{\frac{TR}{T_{10}}} - 1))T_{10}^2, \qquad \mathcal{W}_d = TR(k\sin(\alpha) - S\cos(\alpha)) \tag{4}$$

To minimize the leading order error in T_{10} with respect to the choice of flip angles, we need to minimize the following quantity:

$$\mathcal{W} = \frac{\left[\mathcal{E}_S(e^{\frac{TR}{T_{10}}} - \cos(\alpha)) - \mathcal{E}_k \sin(\alpha)(e^{\frac{TR}{T_{10}}} - 1)\right]}{k\sin(\alpha) - S\cos(\alpha)}. \tag{5}$$

The denominator $F = F(\alpha; TR, T_{10}, k) = k\sin(\alpha) - S\cos(\alpha)$ is an increasing function of the flip angle α for all TR, T_{10} and k. Thus, there exists a constant $C > 0$ such that $F(\alpha; TR, T_{10}, k) > C$ for all α, TR, T_{10} and k. We deduce

$$|\mathcal{W}| \leq \frac{\left|\left[\mathcal{E}_S(e^{\frac{TR}{T_{10}}} - \cos(\alpha)) - \mathcal{E}_k \sin(\alpha)(e^{\frac{TR}{T_{10}}} - 1)\right]\right|}{C}. \tag{6}$$

Therefore $\min_\alpha |\mathcal{W}|$ is less than the minimum of the RHS of (6) with respect to α. Thus, the minimization problem reduces to the minimization of $\mathcal{E} = \|\mathcal{X}_1 - \mathcal{X}_2\|$ with respect to the flip angle α, where $\mathcal{X}_1$ and $\mathcal{X}_2$ denote the random variables

$$\mathcal{X}_1 = \mathcal{E}_S(e^{\frac{TR}{T_{10}}} - \cos(\alpha)), \qquad \mathcal{X}_2 = \mathcal{E}_k \sin(\alpha)(e^{\frac{TR}{T_{10}}} - 1). \tag{7}$$

Assuming Gaussian errors $\mathcal{E}_S \sim N(0, \sigma_S)$ and $\mathcal{E}_k \sim N(0, \sigma_k)$ for every pixel, it follows that $\mathcal{X}_1 \sim N(0, \sigma(\mathcal{X}_1))$ and $\mathcal{X}_2 \sim N(0, \sigma(\mathcal{X}_2))$, where

$$\sigma(\mathcal{X}_1) = \sigma_S(e^{\frac{TR}{T_{10}}} - \cos(\alpha)), \qquad \sigma(\mathcal{X}_2) = \sigma_k \sin(\alpha)(e^{\frac{TR}{T_{10}}} - 1).$$

Because $\mathcal{X}_1$ and $\mathcal{X}_2$ have the same sign (see [1] for proof), the error in T_{10} estimation becomes smallest when these standard deviations are as close to each other as possible for appropriate choice of flip angle. We can therefore obtain the unique optimum flip angle if we choose it such that:

$$e^{-\frac{TR}{T_{10}}}(\sigma_S \cos(\alpha) - \sigma_k \sin(\alpha)) - (\sigma_S - \sigma_k \sin(\alpha)) = 0. \tag{8}$$

This non-linear function of α may be solved using the Gauss-Newton method. We can overcome the spatial constraint that equation (8) imposes by choosing a global (independent of T_{10}) flip angle which minimizes the error in the larger T_{10} values by rearranging equation (8) and taking:

$$\boxed{\frac{e^{-\frac{TR}{T_{10\,\max}}}(\sigma_S \cos(\alpha) - \sigma_k \sin(\alpha))}{(\sigma_S - \sigma_k \sin(\alpha))} = 1.} \tag{9}$$

In doing so, we do not lose accuracy in the optimum flip angle and T_{10} estimation. The reason is that the resulting flip angle does not depend strongly on the exponential factor: for given TR from the pulse sequence protocol, the optimum flip angle as a function of T_{10} in the allowable range for the breast[1] is practically constant, with variations that do not exceed one degree. In general, the flip angle variation is an increasing function of the ratio σ_S/σ_k. Thus the method is also robust for variable T_{10} and the accuracy is guaranteed for the upper range of T_{10} values that are characteristic of tumor types.

3.1 Noise model for k and iterative multiple optimum flip angle determination

The method for determining optimum flip angle acquisition using (9), requires a Gaussian white noise model for the signal S and k. The former can be determined by an off-line phantom experiment. To obtain the value of T_{10} we require also the value of k at each voxel. For those requirements we use the method of Armitage et al. to determine an optimum triplet of angles for T_{10} mapping. The method of Armitage et al. [2] for determining the critical flip angles in a FSPGR echo pulse sequence which minimize the error in T_{10} estimation is summarized as follows: For a fixed value of T_{10} in $\{T_{10}\}$ and for given TR and randomly chosen k, the authors obtain the value of the signal S_i subject to a random flip angle α_i via equation (1). Then they corrupt S_i via Gaussian white noise of a given standard deviation and produce N corrupted signals $\hat{S}_i^N$ which they use to obtain the corresponding corrupted values $\hat{T}_{10i}^N$. For all $\alpha_i \leq 90^o$, they generate a corresponding sequence $a_i(T_{10}) = \{\hat{T}_{10i}^N\}_{\alpha_i}$ for each T_{10}. They choose $\mathcal{F}(\alpha_i) = \sum_{\{T_{10}\}} \sigma(a_i(T_{10}))$ as their objective minimization function ($\sigma(:)$ denoting standard deviation) and obtain three low flip angles $\alpha_1 = 3^o$, $\alpha_2 = 10^o$, $\alpha_3 = 17^o$, each corresponding to a minimum of $\mathcal{F}$. They derive an estimated value for T_{10} by fitting of the three corresponding signal acquisitions following [8].

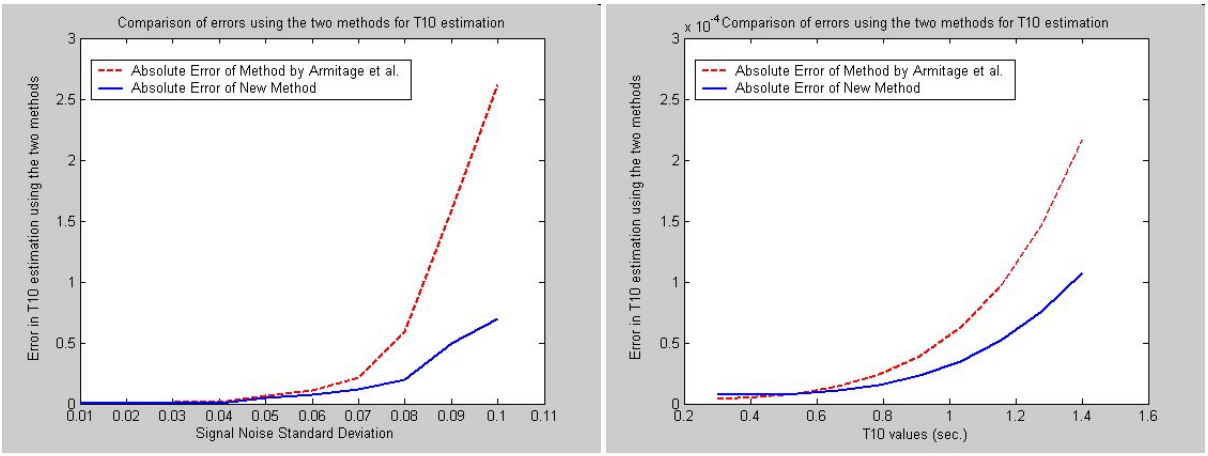
Once the triplet of optimum flip angles is known via the method of Armitage et al. along with the noise model for the signal obtained using a phantom experiment, the Monte-Carlo method can be used to generate a noise model for k as follows: For any random fixed values for k and T_{10}, the given pulse sequence parameters and the three obtained optimum flip angles ([2]), we can generate corresponding signal values. Then we corrupt these signals by the given noise model for the signal and we fit the resulting three signals following [8] in order to obtain the corrupted estimate for k, denoted by $\hat{k}$. Repeating the procedure enough times produces a statistical distribution for $\hat{k}$, whose standard deviation provides the desired noise model for the k map. Once this is known, equation (9) can be used to provide the next optimum flip angle acquisition. Thus, a new flip angle acquisition is obtained from the known three angle acquisitions automatically, at no additional computational cost. Having a fourth optimum flip angle acquisition, we can perform new Monte-Carlo simulations with given noise model for the signal as above to obtain a new, more accurate estimate for the noise model for k. This in turn will result to a new predicted optimum flip angle from equation (9) and the process can be repeated iteratively, until the newly predicted optimum flip angle acquisition is closer to any previously determined optimum flip angle acquisition specified by a given threshold. This is an iterative procedure for computing optimum flip angle acquisitions once the parameters of the scanning sequence are provided and results to more accurately determined T_{10} values according to [7], [2]. All Monte-Carlo simulations and the determination of optimum flip angles can take place off-line once the parameters of the scanning sequence have been provided.

3.2 Comparisons with current protocols for accurate measurement of longitudinal relaxation time and validation on synthetic data

To compare our method with that of Armitage et al. [2] in terms of numerical stability and robustness, we work as follows: for any random value of T_{10} within $\{T_{10}\}$, fixed $TR = 0.0089$sec as chosen in [2] and a random value of k, we evaluate the signals corresponding to the three flip angles predicted by [2] ($3^o, 10^o, 17^o$). We corrupt each of these three signals according to a given Gaussian white noise model with standard deviation σ_S which is fixed from now on. We then fit the signals to compute the estimated T_{10} following [2] and we also predict a Gaussian noise model for k by Monte-Carlo simulations as described in Section 3.1. This determines a fourth optimum flip angle α_4 via equation (9). We then use the given k and T_{10} values to simulate a signal acquisition at the flip angle α_4 and we corrupt the resulting signal by the given noise model. We then fit the noisy data corresponding to the

[1] From now on, we denote the set of allowable T_{10} values for the breast by $\{T_{10}\}$.

acquisitions $\alpha_1 = 3^o$, $\alpha_2 = 10^o$, $\alpha_3 = 17^o$ and α_4 to equation (3) and determine T_{10}. During this process we have produced a statistical distribution for T_{10} predicted values that corresponds to the statistical distributions of the corrupted signals that were used for its derivation. This enables us to define an objective minimization function $f(T_{10}) = \sigma_{T_{10}} \|T_{10} - mean(T_{10})\|$ as a measure of testing the performance of the two algorithms. Figure 1.(a) demonstrates the comparison of the performance of the method of Armitage et al. against the proposed method and plots $f(T_{10}$ against signal noise which ranges from 1% up to 10% of the original signal. Figure 1.(b) compares the two methods using f as a comparison measure, for the whole range of T_{10} values when the signal noise is 1% of the original signal.



(a) (b)

Figure 1. (a) Plot of the objective minimization function $f(T_{10}) = \sigma_{T_{10}} \|T_{10} - mean(T_{10})\|$ for different noise model for the signal (standard deviation) using the two methods ranging from 1% up to 10% of the original MR signal. The superiority of the new method becomes apparent with increasing noise. (b) Plot of $f(T_{10})$ for T_{10} (in sec.) ranging along the allowable values for the breast using the two methods for fixed noise equal 1% of the original signal. This time, the superiority of the new method is evident only for large T_{10} values which are used for classifying tumors.

4 Discussion and Conclusion

This paper presents a new approach in determining multiple optimum angle acquisitions for MR breast cancer diagnosis, using measurements of the T_{10} longitudinal relaxation time which is characteristic of tumor types. The method is based on error analysis of the signal model that enables an analytic formulation of the error in T_{10} estimation, in terms of pulse sequence parameters and noise. This is used to define an optimum angle acquisition via a single equation. The method of Armitage et al. is used as a stepping stone towards estimating the parameters of our equation in order to predict new optimum flip angle acquisitions iteratively, until some pre-defined tolerance has been reached. The analytic formulation of the method clarifies for first time explicitly the role of noise and pulse sequence parameters in selecting optimum flip angles in FSPGR echo sequences. In vitro experiments demonstrated the proposed method to be more stable to signal noise and more robust for all T_{10} values in $\{T_{10}\}$ when compared with the method of Armitage et al [2]. The new protocol is expected to provide highly accurate T_{10} maps and would be a valuable tool towards reliable assessment of breast disease.

References

1. G. Ketsetzis & M. J. Brady. "Optimum flip angle selection in contrast enhanced magnetic resonance imaging.", 2004. Submitted for publication.
2. P. Armitage, C. Behrenbruch, J. Brady et al. "Optimising flip-angles for improved T_1 calculation in 3D contrast-enhanced breast MR imaging." In *Proc. ISMRM*, p. 562. 2001.
3. B. Bone, Z. Pentek, L. Perbeck et al. "Diagnostic accuracy of mammography and contrast-enhanced MR imaging in 238 histologically verified breast lesion." *Acta. Radiol.* **38**, pp. 489–496, 1997.
4. G. Kacl, P. Liu, J. Debatin et al. "Detection of breast cancer with conventional mammography and contrast-enhanced MR imaging." *Eur. Radiol.* **8**, pp. 194–2000, 1998.
5. P. S. Tofts, B. Berkowitz & M. D. Schnall. "Quantitative analysis of dynamic Gd-DTPA enhancement in breast tumors using a permeability model." *Magnetic Resonance in Medicine* **33**, pp. 564–568, 1995.
6. C. Hulka, B. Smith, D. Sgroi et al. "Benign and malignant breast lesions: differentiation with echo-planar MR imaging." *Radiology* **197**, pp. 33–38, 1995.
7. R. Gupta. "A new look at the method of variable nutation angle for the measurement of spin-lattice relaxation times using fourier transform NMR." *Journal of Magnetic Resonance* **25**, pp. 231–235, 1977.
8. R. Venkatesan, W. Lin & E. M. Haacke. "Accurate determination of spin-density and T_1 in the presence of RF-field inhomogeneities and flip-angle miscalibration." *Magnetic Resonance Medicine* **40**, pp. 592–602, 1998.

Generation of multi-phase (4D) CT data based on a nonrigid voxel intensity registration approach and MR breath hold scans

Yassine Bouchareb[a*], Rolf Janka[b] and Willi A. Kalender[a]

[a]Institute of Medical Physics, University of Erlangen-Nuremberg, 91052 Erlangen, Germany, [b]Institute of Diagnostic Radiology, University of Erlangen-Nuremberg, 91054 Erlangen, Germany

Abstract. The aim of the work presented here is to propose an approach to generate multi-phase CT data that make use of a non rigid voxel intensity registration technique and 3D MR breath hold scans. The proposed approach is based on a respiratory motion model derived from 3D MR imaging scans. The 3D MR scans were acquired on 8 patients at different respiration levels over the respiratory cycle. All patients were imaged at four positions between full exhalation and full inhalation while holding their breath for about 15 seconds for each scan. For this purpose, a Turbo flash T1 imaging sequence was used to enhance the signal intensity. All patients had a diagnostic CT scan at full inspiration prior to the MR imaging session on the same day. To quantify the respiratory motion throughout the breathing cycle, all 3D volumes were registered with respect to full expiration using a non rigid voxel intensity-based technique. The resulting sequence of model parameters, reflecting the respiratory motion over the respiratory cycle, was used to generate the multi-phase CT data from the single CT volume. Prior to the generation of the multi-phase CT data, the CT volume was registered to the corresponding breath-hold MR volume. The quality of the registration and fusion results and the generated CT breathing phases was checked by visual inspection and was further assessed by measuring the mismatch between the centres of tumour lesions with well defined borders in both CT and MR. The results obtained by this approach indicated that the mismatch between the centres of lesions between CT and MR was in the order of the slice thickness of CT and MR data, which is suitable for reducing respiration motion artefacts in our intended clinical application: the use of multi-phase CT data for phase-dependent respiratory motion correction in combined PET/CT scans of the chest and the abdomen.

1 Introduction

In recent years there has been a growing interest in acquiring multi-phase or four dimensional (4D) Computed Tomography (CT) data due to the lack of motion information of moving structures in single three Dimensional (3D) CT volumes, essential for chest and abdomen studies. Potential clinical applications could be radiation therapy for quantifying the extent of tumour and recently, with the introduction of the combined PET/CT scanners, the problem of respiratory motion-induced artefacts due to the use of CT images for attenuation correction of the PET emission data. The common requirement for the above-mentioned clinical applications is the availability of 4D CT images in which the internal motion as a function of the respiratory cycle can be quantified. However, and due to dose considerations in CT, the acquisition for multi-phase 4D CT data sets influenced by respiratory motion is not allowed in current clinical routine practice. Different respiratory gating techniques, retrospective image reconstruction and scanning protocols using both single slice and multi-slice CT scanners have been investigated [1-3], however they remain complex and impractical in patient-by-patient basis. The aim of this work is to address this issue by proposing an image processing-based approach to generate the multi-phase CT data that makes use of a fully automatic non rigid voxel intensity registration technique and breath hold Magnetic Resonance (MR) scans.

2 Method and Materials

2.1 Data acquisition

Both CT and MR scans were acquired on eight patients (6 male and 2 female; mean age 59 years, range 49-79 years) on the same day. The patients had cystic lesions mainly in the liver or the kidneys. All patients were first scanned in a 16-slice MSCT scanner (Sensation 16, Siemens Medical Solutions, Erlangen, Germany), according to routine protocols at maximum inspiration using intravenous contrast agent. The scans were performed with elevated arms in breath-hold inspiration. Image reconstruction resulted in 512 x 512 images with 0.72 x 0.72 mm^2 pixel size and a slice thickness of 5 mm using a 5 mm reconstruction increment.

The 3D MR scans were carried out on a 1.5 T scanner (Magnetom Symphony, Siemens, Erlangen, Germany). All patients were imaged at four positions between full exhalation and full inhalation while holding their breath for about 15 seconds for each scan. For this purpose, a Turbo FLASH T1weighted imaging sequence was used. A

* yassine.bouchareb@imp.uni-erlangen.de

chest belt was used to monitor the quality of each breath hold. The volume covered a 380 mm x 260 mm field of view with 52 transverse slices of 5 mm slice thickness. The reconstructed images were 512 x 384 with an image resolution of 1.34 x 1.34 x 5.00 mm.

2.2 Registration and motion quantification

To quantify the respiratory motion throughout the breathing cycle, all 3D MR volumes were registered with respect to full-expiration using a non rigid voxel intensity based technique. The resulting sequence of model parameters, reflecting the respiratory motion over the respiratory cycle, was used to generate the multi-phase CT data from the single CT volume. Prior to the generation of the multi-phase CT data, the CT volume was registered to the corresponding breath-hold MR volume. Figure 1 illustrates the process of generating the 4D CT data using the estimated motion parameters from the MR breath-hold image volumes.

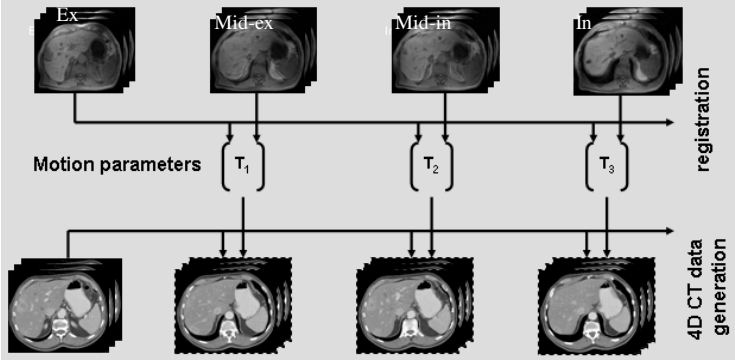


Figure 1. Sketch of respiratory motion quantification using the non rigid registration for 4D CT data generation. Top row are the acquired breath hold MR volumes; bottom left is the acquired CT volume; on its right (dashed borders) are the possible generated CT volumes at the corresponding MR respiratory phase.

The key ideas of our non rigid voxel intensity based registration technique consist of considering a hierarchical transformation model combined with a multiresolution, as widely used in previous studies [4,5] to speed up convergence and avoid local minima, and a multi-optimisation strategy. The idea of using a multi-optimisation scheme as described first in Bouchareb et al. [6] is to take advantage from the convergence behaviour of non-derivative optimisation methods when aligning image sets realised in extreme differences in terms of breathing protocols during data acquisition. It was validated using FDG-PET and CT phantom and patient data [6].

The Normalised Mutual Information (NMI) proposed in [7] is an invariant similarity measure with respect to the volume of overlap between two images and it was used in this study. It was derived from the Mutual Information (MI) similarity measure, introduced in image registration by Viola [8]. NMI of two data volumes A and B, physically conveys the amount of information that A contains about B, or *vice versa*. It can be defined in terms of the marginal entropies $H(A)$ and $H(B)$ of the images, combined with their joint entropy $H(A,B)$, as follows:

$$NMI(A,B) = \frac{H(A) + H(B)}{H(A,B)} \qquad (1)$$

The transformation mode used in our formulation of the problem is a generalized affine transformation (T), considered as cumulative effect of a series of scaling (S), shearing (H), rotation (R), and translation (D). The order of combining these transformations has been restricted to the following: $T = D \times R \times H \times S$.

The expanded homogeneous 4 x 4 affine transformation appears as

$$T = \begin{vmatrix} e_{xx} & e_{xy} & e_{xz} & d_x \\ e_{yx} & e_{yy} & e_{yz} & d_y \\ e_{zx} & e_{zy} & e_{zz} & d_z \\ 0 & 0 & 0 & 1 \end{vmatrix} \qquad (2)$$

Where $\{d_x, d_y, d_z\}$ is the 3D translation vector and the e_{ij} elements of the 3x3 submatrix encompass the combined effect of 3D rotations $\{r_x, r_y, r_z\}$, scalings $\{s_x, s_y, s_z\}$, and shearings $\{\theta_{xy}, \theta_{yz}, \theta_{zx}\}$ transformation parameters.

The desired solution of our registration problem while maximising NMI, equation (1), is formulated as follows:

$$T^* = \arg\max_{T \in S_T} NMI(A, T(B)) \qquad (3)$$

The major component in the overall registration process is the optimisation method. This optimisation method is called repeatedly and so should be efficient with respect to the resolution of the images and the similarity measure used to align these images. In the present work, to estimate the transformation T^*, equation (3), a multi-optimisation scheme is considered, in which a different optimisation method is used at each resolution level. Powell's method is used at lower resolution levels and the Nelder-Mead simplex optimisation is used at the full resolution level. Being very sensitive to initialisation, the simplex method is initialised with the optimal solution found by Powell's method at the previous resolution level. The best order in which the parameters have been optimised was investigated at the low resolution level. In our study, it was found that optimising the translations first or the rotations first depending on which movement is more significant speed up the registration process. This is due the fact that the transformation matrix, equation (2), is better conditioned under a given optimisation order. A detailed diagram of the Multi-Resolution Multi-Optimisation (MRMO) algorithm is shown in figure 2.

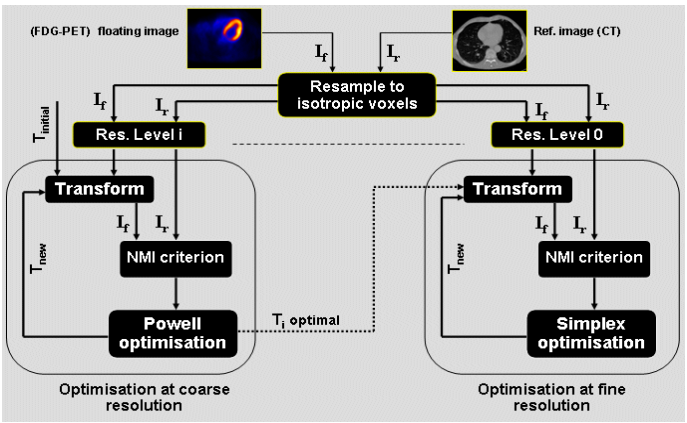


Figure 2. Diagram of the MRMO algorithm using a hierarchical transformation model and NMI. The two volumes are first resampled to isotropic voxels; then the resolution levels (RL) are created (3 RL were used here); Powell's optimisation is used at the coarse resolution and Simplex optimisation at the fine resolution level while initialized with the optimal transformation found by Powell's method.

3 Results

The evaluation of the quality of the registration and fusion results and the generated CT breathing phases was checked by visual inspection and was further assessed by measuring the mismatch between the centres of tumour lesions (1 to 2 lesions localized in the liver or kidneys per patient) with well defined borders in both CT and MR. The visual assessment and all measurements were made using the 3D viewing application integrated in the cross-modality *Syngo*® software (Siemens) for all breathing phases. Figure 3 shows some fused CT and MR images with cystic lesions.

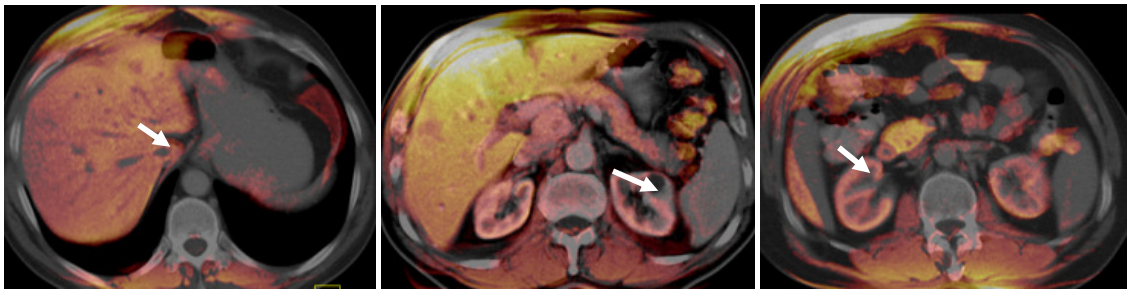


Figure 3. Axial fused CT and MR images from three patients with cystic lesions. As pointed out by the arrows; left: cystic lesion in the liver; middle: cystic lesion in the left kidney; and right: cystic lesion in the right kidney

The visual assessment of the fusion results indicated the successful alignment of CT and MR image sets in 7 from the 8 patients included in this study. The analysis of all measurements given in table 1 were made on the fused images for all subjects showed that the mismatch between the centres of lesions was ranging from 0.8 to 1.6 mm

at full expiration and 4.1 to 6.2 mm at full inspiration with mismatches up to 9 mm for one patient; probably due to the positioning or voluntary movement of this patient. In most cases, the best match; i.e. the smallest mismatch is obtained at the full expiration phase, followed by mid-expiration, mid-inspiration and then full inspiration.

Subject	Full-inhale	Mid-inhale	Mid-exhale	Full-exhale
P1	4.1	3.3	2.1	1.2
P2	3.0	2.5	1.6	0.8
P3	3.2	1.8	1.8	1.2
P4	6.5	4.2	3.8	1.3
P5	9.0	6.0	4.3	2.4
P6	3.3	2.0	3.1	1.1
P7	5.4	1.8	2.0	1.6
P8	7.2	5.5	3.2	1.5

Table 1. Mismatch between the centres of lesions in CT and MR (in mm) measured for all four breathing phases

These measurements confirm the higher reproducibility of the full expiration phase compared to full inspiration as reported in previous studies. For 7 subjects and over all breathing phases, the mismatches were in the order of the slice thickness in CT and MR, validating our approach in generating multi-phase or the 4D CT data.

4 Conclusions and future work

The potential of the developed approach in generating 4D CT data based on MR breath hold scans and using a fully automatic non rigid registration technique, validated previously on PET and CT phantom and clinical data has been demonstrated in 8 patients. The 4D CT images contain respiratory motion information not available in the single CT volume, which helps to understand the motion of the tumour and surrounding tissues, essential to plan radiation therapy, for example. Although the derived motion model allows accuracy in the order of the slice thickness according to our experiments, the registration technique; emphasing the importance of using a multi-optimisation strategy in combination with the multi-resolution scheme, which has been used to quantify the motion information, is limited to global movement in the chest and abdomen. Therefore, further work will focus on including free-form deformations to account for local deformation of different organs in the chest and abdomen on which the quantification of the motion will be based and thereby the generation of the 4D CT data. Ongoing investigations concern the use of this multi-phase 4D CT data for phase-dependent attenuation correction of the PET emission data in combined PET/CT scans to minimise the respiratory-induced motion artefacts for chest and abdomen tumour studies.

Acknowledgements

We would like to thank the Bavarian Science Foundation for funding this project (grant number AZ 460/01).

References

1. T. Pan, T.Y. Lee, E. Rietzel, et al. “4D-CT imaging of volume influenced by respiratory motion on multi-slice CT”. *Med. Phys*. **31 (2)**, pp. 333-340, 2004.
2. T. Pan, K. Acharaya, S. Kohlmyer et al. “Development of a 4D respiratory gating system for PET-CT”. *SNM 50th Annual Meeting*, New Orleans, June 2003.
3. M. Kachelriess & W.A. Kalender. “Electrocardiogram-correlated image reconstruction from sub second spiral computed tomography scans of the heart”. *Med. Phys*. **25**, pp. 2471-2431, 1998.
4. C. Studholme, D.L.G. Hill & D. Hawkes. “ Multi-resolution similarity measures for MR-PET registration”, *in Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Poala, Brest, France, Eds. Dordrecht: Kluwer Academic Publishers, pp. 287-298, 1995.
5. J.P.W. Pluim, J.B.A. Maintz & M.A. Viergever. “Mutual information matching in multiresolution contexts”. *Image and Vision Computing*. **19**, pp. 45-52, 2001.
6. Y. Bouchareb, W. Römer, G. Platsch et al. “Normalised mutual information-based nonrigid registration (fusion) approach of FDG-PET and CT scans”. *RSNA Annual Meeting*, Chicago, Dec. 2003.
7. C. Studholme, D. Hill & D. Hawkes. “An overlap invariant entropy measure of 3D medical image alignment”. *Pattern Recognition* **32 (1)**, pp. 71-86, 1999.
8. P. Viola. *Alignment by maximization of mutual information*. Ph.D. thesis, MIT, Cambridge, Ma, 1995.

3D MR liver modelling for use in pre-surgical planning

H Ibrahim[a]*, M Petrou[a], K Wells[a], S Doran[a] and O Olsen[b]

[a]CVSSP, School of Electronics and Physical Sciences, University of Surrey, Guildford, GU2 7XH, UK,
[b]Radiology Department, Great Ormond Street Hospital for Children NHS Trust, London, WC1N 3JH, UK

Abstract. Three dimensional (3D) visualization has the potential to significantly ease the decision making process in pre-surgical planning. The first stage of creating a 3D model for this purpose is to segment the liver from MRI data. We propose to automate a segmentation technique known as intelligent scissor to segment the liver. The user only needs to select an initial slice, and the method is executed automatically. From the initial slice, the contour propagates inside the volume and segments the liver in every slice in the dataset. Thus automatic intelligent scissors estimate the seed points automatically and connect these points based on a dynamic programming algorithm. Our results show that the combination of segmented outputs from orthogonal directions can improve the performance of the segmentation compared with the use of slices along a single orientation only.

1 Introduction

The extent of hepatic disease (e.g. cirrhosis or metastatic cancer) can be assessed by MRI. In pre-surgical planning, contiguous 2D slices are viewed by the clinician. However, a key limitation with imaging is that contrast-enhanced blood vessels may be confused with small tumours. This problem can be obviated by viewing a 3D model of the liver anatomy and its associated vasculature, and any pathology.

Segmentation is the first step in the creation of such a model. Given that manual segmentation is impractical, we propose an automatic method for segmenting the hepatic volume as a first stage in model creation. Unlike related work in CT [1] the MR target volume cannot be globally windowed or segmented using a global threshold. This is because abdominal MR data contains many objects with similar greyscale intensities, which may also be in contact with one another. Nonetheless, edge information is still present and for this reason, we have utilised a technique called intelligent scissors, as originally proposed by Mortensen and Barrett [2], which is a semi-automatic contour-based segmentation tool. Intelligent scissors can be considered as a combination of edge-based and dynamic programming segmentation. In its usual form, this technique requires a human operator to determine the starting and termination points for determining the desired edge segments, interactively. The resultant edge segment, which connects two defined points, is created based on the lowest cost path according to Dijkstra's algorithm [3], utilising a cost function composed of gradient and zero-crossing information.

In this work, we propose a method to automate the segmentation of the liver using intelligent scissors. In this method, the user only need select one axial slice as the starting slice; the ensuing segmentation process is executed fully automatically. We also propose a technique to combine several segmentation results in order to produce a more reliable output.

2 Method

2.1 Pre-processing

We pre-process the 3D MR input data to reduce the artefacts, and enhance its quality. First, to reduce the artefact that is due to the sensitivity of the RF coil(s) used in the MR scanner, we locally enhanced the dataset, in 3D space, based on the technique described in [4]. Then, to reduce the additive noise, we filter the data with a cubic median filter of size 7 x 7 x 7. Next, to re-enhance the edges, we implement a toboggan enhancement, which is based on [5], but extended to 3D.

To emphasise the liver outline, we decided to use a method that salienates features in the edge map which have some perceptual meaning. To achieve that, we implemented a pre-attentive segmentation model, proposed recently by Li [6], for the function of the human brain that is responsible for feature salienation. We take the output of this neural-network implementation, after a few iterations, as the pre-attentive segmentation result. When applied to our data, we found that the edge salienation model converges after typically two iterations.

*Email: h.ibrahim@eim.surrey.ac.uk

2.2 Automatic Intelligent Scissors

The automation of the intelligent scissors approach segments the liver on a contiguous slice-by-slice basis, but using the *a-priori* information from the previous segmentation slice. The main idea of our method is to propagate a contour from an initial slice to the rest of the slices in the dataset.

Selection of the initial slice We select one axial slice from the input data as the initial slice. This initial slice is selected based on visual inspection. The criteria for this selection are: (1) the liver has a good contrast with its surroundings, and (2) the shape of the liver preferably has some resemblance to a circle or an ellipse.

Estimation of the seed points in the initial slice Following slice selection, we threshold the corresponding salienation map of the initial slice, and skeletonize the output to get one pixel width edges. In order to estimate the sample points, we first estimate the centre of the curve we wish to identify. We select randomly a triplet of edgels from the edge image, and assume that these points have coordinates (x_1 , y_1), (x_2 , y_2), and (x_3 , y_3). The coordinates of the centre of the circle they define, (x_c , y_c), are given by the following equations:

$$x_c = \frac{(y_3 - y_1)(y_3 - y_2)(y_2 - y_1) + (x_3^2 - x_3^2)(y_2 - y_1) - (x_2^2 - x_1^2)(y_3 - y_2)}{2[(x_1 - x_2)(y_3 - y_2) - (x_2 - x_3)(y_2 - y_1)]} \quad (1)$$

$$y_c = \frac{(x_3 - x_1)(x_3 - x_2)(x_2 - x_1) + (y_3^2 - y_3^2)(x_2 - x_1) - (y_2^2 - y_1^2)(x_3 - x_2)}{2[(y_1 - y_2)(x_3 - x_2) - (y_2 - y_3)(x_2 - x_1)]} \quad (2)$$

To estimate the best centre from the fragments of the contour of the liver, we use a random selection of 3N sets of unique triplets, where N is the number of the edgels in the image. We create a 2D accumulator array, which represents a 2D histogram of the (x_c , y_c) values computed from each triplet. From this accumulator array, we select the position of the peak of the 2D histogram. This position represents the ideal centre of the edgel fragments we have. We project straight lines every 15^o emanating from (x_c , y_c). Then, we identify as seed points the first intersection points these lines have with the edge fragments. The desired edges of the target object between these seed points are refined by invoking intelligent scissors, and using these seed points as start and stop points that would otherwise be manually defined. Repeated application of this procedure at each 15^o segment produces the initial estimate of the liver boundary in the initial slice.

Propagation of the contour to the next slice We now assume that the shape of the liver does not change significantly in the next adjacent slice. Thus, we use the above segmented contour as an initial reference contour to segment the next slice. We take the inflection points of the contour as the seed points. These inflection points are calculated by using a technique proposed by Rosenfeld and Johnston [7]. To improve accuracy and save computation time, when the intelligent scissors algorithm is invoked, the cost function is only calculated around the reference contour locally. We estimate this local region using morphological operations, where the local region is defined as the reference contour, dilated 15 times using a 3 x 3 structuring element. This contour propagation approach is then repeated at each subsequent slice in the volume.

2.3 Combination of axial-coronal and axial-sagittal segmentations

The above segmentation result actually defines an initial estimate of the volume of the liver. To refine this re-estimate, we slice this segmented volume in a perpendicular orientation, and thus we get a representation of the liver area in this new direction. Based on this area, we create a reference contour to segment the corresponding slice in this new orthogonal direction. Using the technique described above, we again segment the liver using intelligent scissors.

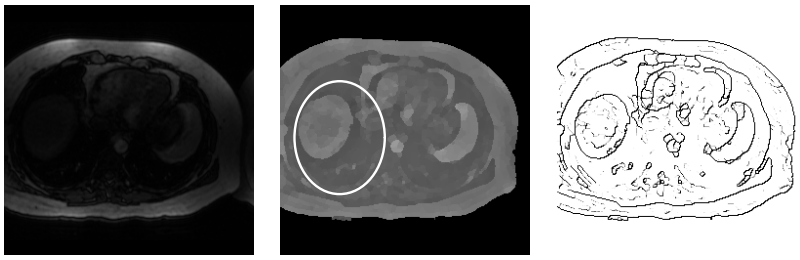
To initiate this process, we use the segmentation result re-oriented from axial direction as the reference of segmentation in the coronal direction. We refer to this as the axial-coronal segmentation. Similarly, the result from the axial direction is also used as the reference to segment the liver in the sagittal direction. We call this the axial-sagittal segmentation. We consider the volume defined by the axial-coronal segmentation as V_{ac}, and the volume defined by the axial-sagittal segmentation as V_{as}. The combined output volume, V_{out}, is then given by:

$$V_{out} = V_{ac} \bigcap V_{as} \quad (3)$$

which means that we consider that a voxel is valid as a liver voxel only when it is detected in both orientations.

3 Results and Discussion

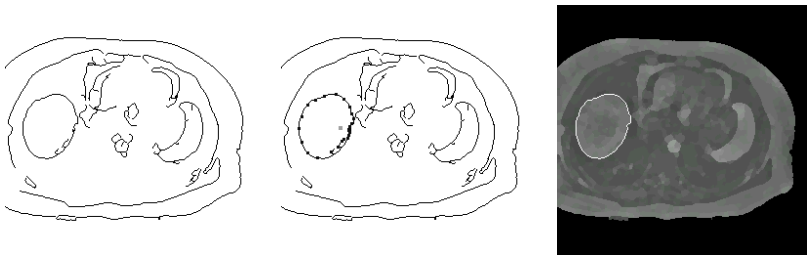
Figure 1 shows one axial slice from a 3D MR input dataset used in our experiment. It is shown that the liver can be perceived relatively easily in the pre-processed image compared to its original input. This suggests that pre-processing improves the quality of the input data.



(a) (b) (c)

Figure 1. One axial slice from 3D MR input data. (a) Input image. (b) Image (a) after pre-processing. (c) The salienation map of image (b). Object in circle in (b) represents liver segment in the initial target slice.

The slice shown in Figure 1 fulfils the requirement for the initial slice, thus we take this slice as our initial slice. Figure 2 shows each step taken to segment the liver boundary from this slice.



(a) (b) (c)

Figure 2. (a) Image from Figure 1(c) after thresholding and skeletonization. (b) 24 sample points detected by projecting straight lines at every 15^o from the centre (indicated by a grey box). (c) The result of intelligent scissors segmentation, using the seed points shown in (b), to delineate the liver.

This final detected contour becomes the reference contour to segment the next neighbouring slice. The contour propagates inside the volume in the axial direction, until the liver in all slices has been segmented. The shaded surface representation of the result is shown in Figure 3(b). This segmented volume is then used as the reference for segmenting the data in the coronal and sagittal direction respectively. The result, which is based on equation (3), is shown in Figure 3(c).



(a) (b) (c)

Figure 3. (a) 3D shaded surface model of the ground truth, which has been segmented manually. (b) 3D shaded surface model of the segmentation in the axial direction. (c) 3D shaded surface of V_{out}.

Figure 4 shows the measurements of error of both segmentations in the axial direction, and V_{out}, relative to the ground truth. By considering N_T as the liver volume in the segmented output, N_G as the liver volume in the ground truth, N_P as the number of voxels detected in the output but not in the ground truth, and N_N as the number

of the voxels detected in the ground truth but not in the output, we define the over detection error, *Eo*, as N_p/N_T, the under detection error, *Eu*, as N_N/N_G, and total detection error, *Et*, as *Eo*+*Eu*.

From this figure, we can see that V_{out} has slightly higher under detection error compared with the segmentation in the axial direction alone. However, the total detection error of V_{out} is much lower. Thus we suggest that the combination of segmentation results using equation (3) produces a better output.

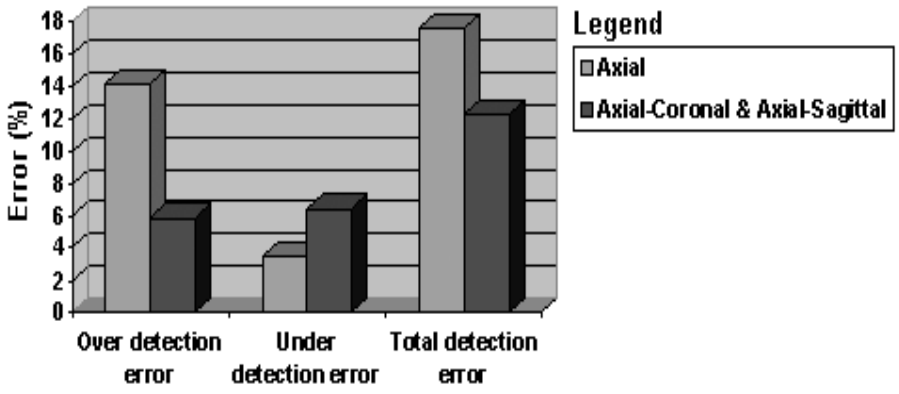


Figure 4. A bar graphs showing the error measurements for the segmentation results in the axial direction, and the combination of axial-coronal with axial-sagittal segmentations, V_{out}.

4 Conclusions

Combining orthogonal segmentation results can reduce the number of voxels that have been misclassified as liver parenchyma. Although under detection error from this technique is slightly increased, over detection error is significantly reduced. Thus, in overall, this combination improves the quality of the segmentation.

References

1. L. Gao, D. G. Heath, B.S. Kuszyk & E. K. Fishman. "Automatic liver segmentation techniques for three-dimensional visualization of CT data." Radiology **2(201),** pp. 359-364, 1996.
2. E. N. Mortensen & W. A. Barrett. "Interactive segmentation with intelligent scissors." *Graphical Models and Image Processing* **60**, pp. 349-384, 1998.
3. D. Bertsekas and R. Gallager. *Data Networks*. Prentice-Hall International Editions, 1987.
4. P. M. Narendra & R. C. Fitch. "Real-time adaptive contrast enhancement." *IEEE Trans. Pattern Anal. Machine Intell.* **PAMI-3**, pp. 655-661, 1981.
5. J. Fairfield. "Toboggan contrast enhancement." *SPIE Applications of Artificial Intelligence X: Machine Vision and Robotic* **1708**, pp. 282-289, 1992.
6. Z. Li. "Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex." *Network: Computation in Neural Systems* **10**, pp. 187-212, 1999.
7. A. Rosenfeld & E. Johnston. "Angle detection on digital curves". *IEEE Transactions on Computers* **22**, pp. 875-878, 1973.

Dynamic MRI with regular under-sampling for increased time resolution: x-f choice

S. Malik, S. Schmitz, J.V. Hajnal

Imaging Sciences Department, Clinical Sciences Centre,
Hammersmith Hospital Campus of Imperial College London

1 Introduction

Image acquisition speed is frequently a limiting factor for dynamic MRI studies, e.g. cardiac imaging and dynamic contrast enhanced MR angiography (DC-MRA). During dynamic image acquisition, k-space is sampled repeatedly in time (t) forming a k-t data space. Temporal resolution is often limited by the time taken to fully sample k space for each new image. To increase temporal resolution sub-sampling schemes can be adopted. Methods for fMRI and cardiac imaging based on regular sub-sampling have been proposed by Madore et al (UNFOLD) [1] and Tsao et al (k-t BLAST and k-t SENSE) [2]. These approaches can be fitted into a general framework for working with k-t space introduced by Xiang and Henkelman [3]. Regular under-sampling in k-t space leads to aliasing in a prescribed way in the conjugate image and temporal frequency domain (x-f space). In the UNFOLD technique the resulting ambiguity is resolved for limited speed-up factors by use of a mask that excludes signals aliased to higher temporal frequencies. Other approaches such as Sliding window rely on filtering to suppress aliases. In k-t BLAST and k-t SENSE, the filtering is adaptively specified for each x-f location based on training data. In this study we have investigated strategies for accelerating DC MRA acquisitions.

DC-MRA image series' contain highly localised dynamic regions, the vessels, surrounded by other tissue, which is relatively static. Figure 1 shows an x-f space representation of a particular line in the frequency encode direction of some real DC-MRA data. In order to gain temporal resolution regular under-sampling is undertaken. To simulate the effect of this, the data shown in Figure 1 has been four-fold under-sampled post acquisition using IDL, generating figure 2. To recover intelligible images one must resolve the ambiguity introduced by aliasing, which mixes signals that would be separate in a fully sampled case.

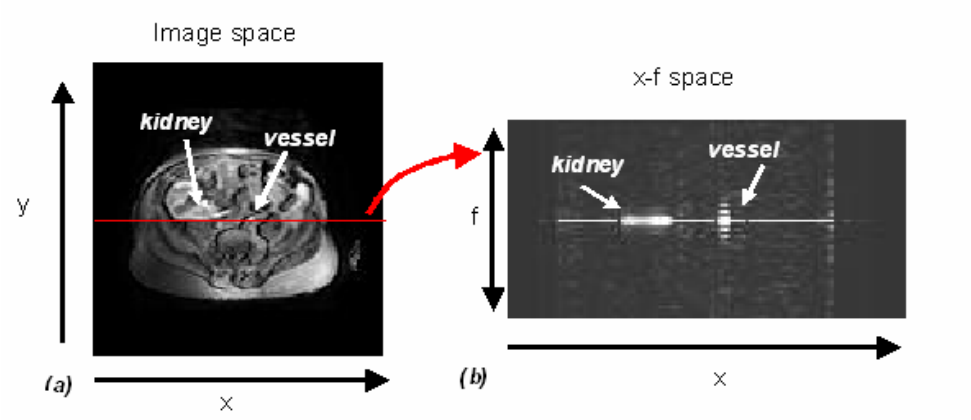


Figure 1. **a)** One time frame from a DC-MRA study. The 'x' direction here is the PE direction, 'y' being the FE direction. Each line for a given y has a separate x-f space. **b)** This x-f space is from the line indicated in part a). A kidney and a blood vessel are apparent in the image, they are distinctly recognisable in the frequency representation. The kidney has a slow rise time from the contrast agent, hence the low frequency response in x-f, whereas the vessel's sharp rise gives the oscillatory response.

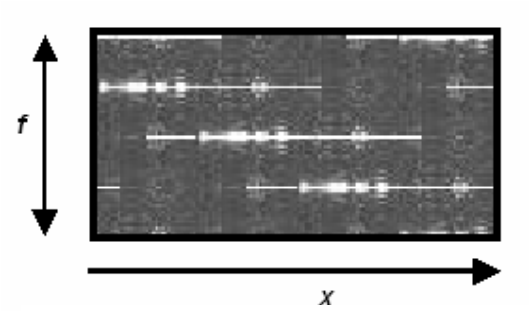


Figure 2. Four fold aliased x-f spectrum from patient data DC-MRA model. The aliases are offset in frequency and in space

2 Methods

Aliasing mixes signals from different locations in the fully sampled x-f space. We note that for a given x, the f spectra generally have high signals near f=0 with signal dropping off rapidly as the modulus of f (|f|) increases. Thus the mixed signals are usually very uneven in intensity, it is more faithful to the data to retain only the dominant alias (i.e. making an x-f choice), rather than to partition the signal. This is illustrated by a four-fold aliased spectrum in figure 3, where contributions from different spatial locations overlap the temporal frequency profile centred at f=0. The alias centred on f=0 is referred to as the 'base band'.

However as figure 4 shows, the true balance of aliased signals is often obscured by noise, which may be large in comparison to the required data when |f| is large. To decide which alias is dominant at any given x-f location we fit the frequency spectrum at a given x with an envelope function based on prior

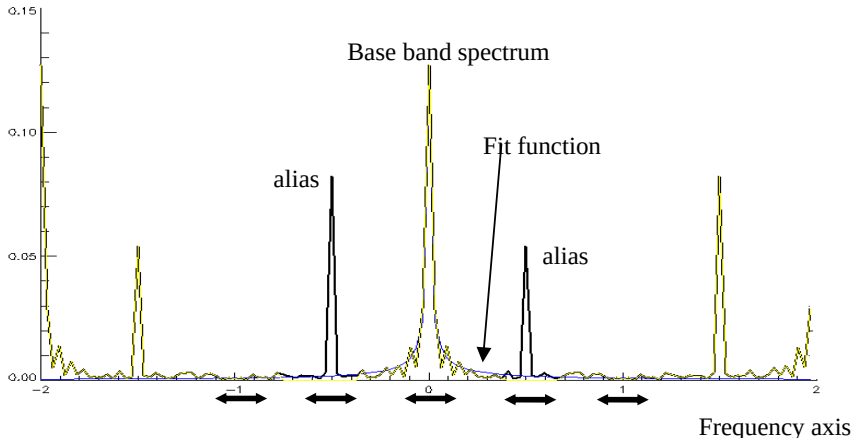


Figure 3. The temporal spectrum from a given 'x' location. There are contributions from other spatial locations centred at non zero frequencies. The alias centred on f=0 is the correct one for this spatial location and is referred to as the 'base band'. Fitting functions to each alias allows their width to be estimated so that they can be removed. Note that in this plot the Nyquist frequency has value 1 on the frequency axis. The absence of an alias at the Nyquist frequency is by chance - there was nothing at the corresponding spatial location in this case.

knowledge. The fitting process exploits high SNR regions of x-f space to allow signal estimates to be made even in regions of low SNR. In the case of DC-MRA the contrast agent causes signals to increase over time and this can be used to define the model function used to fit the frequency spectra.

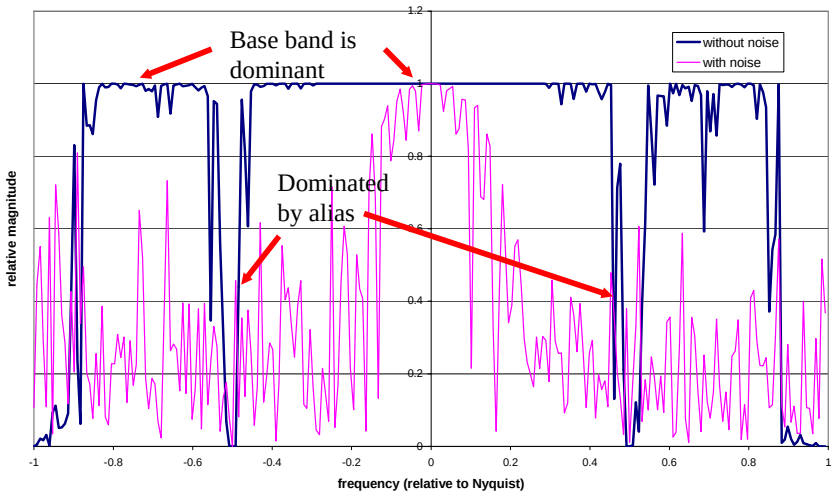


Figure 4. The plots show the square modulus of the base band alias compared to the sum of all contributions. Both plots were taken from the same spectrum. When there is no noise it is clear that the signal is dominated in the main by just one alias, be it the base band so that the relative magnitude is 1, or another alias giving zero. When noise is added this distinction is less clear.

Where the signals are dominated by contributions that are not the required base band signals some form of data processing is required in order to produce final un-ambiguous images. The k-t-BLAST method uses fully sampled (training) data to determine weights that can be used to attenuate the undesired aliases. Sliding window reconstruction works directly on the raw k-t data to estimate the missing points by interpolation, so reducing the intensity of the aliases. In the x-f choice method, wherever base band alias is not dominant this region of x-f space is 'cut out' and replaced with zeros. Because the pattern of aliases is offset in x-f space (figure 2), for any given x location, the spectrum at –f is often dominated by the base band, even when the value at +f is in need of correction. In these cases it is possible to use conjugate symmetry of the frequency spectrum to partially fill in missing data.

The method has been tested on simulations using the IDL environment and on single slice *in vivo* data acquired on a Phillips Intera 3T system. The *in vivo* data was acquired following injection with Gd-DTPA prior to a standard DC-MRA examination. Model data was created by producing a time course of 2-d images (see figure 5) with three circular regions specified as (A) an artery characterised by early, rapid signal increase and (B, C) veins characterised by later slower increases. The fully sampled time series for both the simulated and the in vivo data were decimated to create under-sampled data sets.

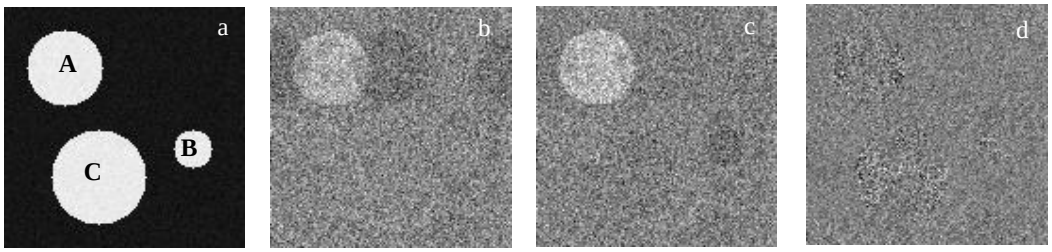


Figure 5. a) A single time point during rapid arterial signal change from the DC-MRA model (A = artery, B, C = veins). **b, c & d)** are all subtractions of reconstructions from four-fold under-sampled data using: sliding window (b), k-t BLAST (c) and x-f choice (d) minus the original. All are at the same time point and have the same window and level. K-t BLAST training data was the central 25% of k-space, the noise level in this data critically effects performance. The x-f choice reconstruction shows no obvious temporal blurring.

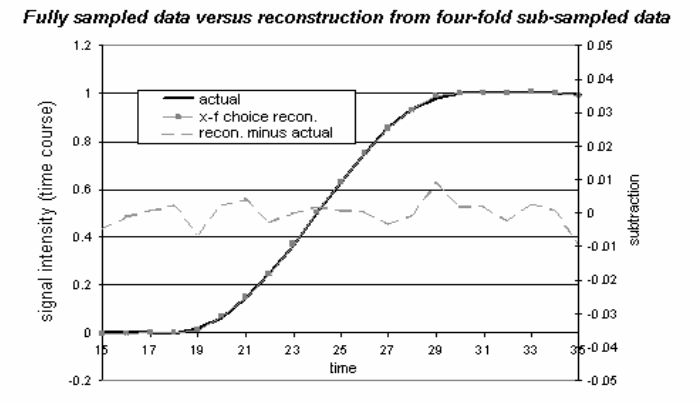


Figure 6. X-f choice does not cause temporal blur in this time series, taken from an ROI in vessel C from fig 5a.

3 Results

A fully sampled simulated DC-MRA single slice data set was under-sampled by a factor of 4 and reconstructed using x-f choice. Subtractions show that with x-f choice errors appear as an increase in incoherent noise over dynamic regions (figure 5d).

K-t BLAST and sliding window reconstructions demonstrate temporal blurring manifesting as coherent error over the dynamic region in the subtraction (figures 5b and 5c). None of the methods demonstrate any obvious spatial blurring, owing to the spatially static nature of the model.

The x-f choice reconstruction has the lowest level of average error, an ROI taken in the lower left vessel (C) displays the time series and subtraction showing no temporal blurring (figure 6).

4 Conclusions

DC-MRA usually yields predictable spectral shapes, making the method ideal to use without training data. Other dynamic studies such as cardiac imaging may also be suitable as potential applications. These may proceed by using training data to calculate line shapes. It would also be possible to combine this approach with existing parallel imaging methods using multiple coils to achieve larger acceleration.

The maximum acceleration factor possible depends on the data itself. The advantage of x-f choice over related reconstruction methods is that high frequency detail can be preserved, which is shown by the lack of temporal blurring.

Acknowledgements

Thanks to Philips Medical Systems for research grant support. SM thanks the MRC for a studentship.

References

1. Madore, B., Glover, G., Pelc, N., 1999. Unaliasing by Fourier-Encoding the Overlaps Using the Temporal Dimension (UNFOLD), Applied to Cardiac Imaging and fMRI. Magn Reson Med 42, 813-828.

2. Tsao, J., Boesinger, P., Pruessmann, K., 2003. k-t BLAST and k-t SENSE: Dynamic MRI with high frame rate exploiting spatiotemporal correlations. Magn Reson Med 50, 1031-1042.

3. Xiang, Q-S., Henkelman, M., 1993. K-Space description for MR Imaging of dynamic objects. Magn Reson Med 29, 422-438

An Improved ASM Approach Using Mixture Models to Prostate Segmentation from MR Images

Yanong Zhu[a*], Stuart Williams[b] and Reyer Zwiggelaar[c†]

[a]School of Computing Sciences, University of East Anglia, Norwich, UK
[b]Department of Radiology, Norfolk and Norwich University Hospital, Norwich, UK
[c]Department of Computer Science, University of Wales, Aberystwyth, UK

Abstract. In Active Shape Models (ASM) methodology, grey-level profiles at landmarks are modelled as a Gaussian distribution. The Mahalanobis distance from a sample profile to the model mean is used to locate the best position of a given landmark during ASM search. We present an improved ASM methodology, in which the profiles are modelled as a mixture of Gaussians, and the probability that a sample is from the distribution is calculated using the probability density function (pdf) of the mixture model. Both improved and original ASM methods were evaluated on synthetic and applied to segmentation of prostate from MR images. The performance comparison demonstrates that the improved ASM method is more generic and shows higher robustness than the original approach.

1 Introduction

The use of Active Shape Models (ASM) has been shown to be an efficient approach to the image interpretation and pattern recognition [1]. Image segmentation using ASM can be divided into two stages. Firstly, a parameterized statistical shape model is built from labelled training images. The grey-level profiles normal to the object boundary through each landmark are extracted and modelled as a single Gaussian distribution. Subsequently, a shape instance is deformed in accordance with the model to search for a boundary which optimally segments the object. This is performed by two key steps: (a) looking for better positions for each individual landmark using Mahalanobis distance and, (b) updating the model parameters to best fit the shape instance to the newly found landmarks. Both steps are crucial to the final search results.

Despite the successful cases of applying ASM, the ASM method still has some limitations and hence can be improved in several ways. Development has been done to improve the shape modelling and generation aspects in ASM method. Cootes et al. [2] used the Gaussian mixtures to model the distribution of the landmarks and hence the shape variation. Rogers and Graham [3] improved ASM by using M-estimator and random sampling approaches to robust parameter estimation instead of Gaussian distribution based estimation. Twining et al. [4] described the use of Kernel Principal Component Analysis (KPCA) to model the variability in a class of shapes. On the other hand, van Ginneken et al. [5] used optimal image features instead of grey-level profiles for ASM search, and applied a *k*-Nearest Neighbour (*k*NN) classifier to find the displacement of landmarks. These improvements have achieved credible performance and largely increased the efficiency and robustness of ASM methods.

We have applied ASM to segment objects of interest from image data sets. Some of our experiments showed that, when the grey-level variations around the object border are more complicated than a single Gaussian distribution, the use of Mahalanobis distance to measure the distance from a sample to the mean of the distribution becomes inaccurate and consequently causes invalid search results. In ASM, the derivative profiles are used to reduce the effects of global intensity changes, which only represent the relative intensity change along the profiles.

In this paper, we concentrate on the modelling aspects of grey-level profiles, as well as locating of landmark positions using the profile models. To overcome the ASM limitations mentioned above, instead of a single Gaussian distribution and Mahalanobis distance, we have used a Gaussian mixture to model the profiles, and, the probability that a sample profile comes from the distribution is measured by the total probability of sub-distributions. We would like to emphasis that this distinct from Cootes et al. [2] where Gaussian mixtures were used to model the shape variation but not the profile variation.

*Email: yz@cmp.uea.ac.uk
†Email: rrz@aber.ac.uk
This project is funded by Prostate Research Campaign UK.

2 Method

In our improved ASM method, the grey-level profiles are treated as a Gaussian mixture model. We assume that the sum of a certain number of Gaussian distributions can represent the distribution of these profiles. With this optimal Gaussian mixture, the probability that a sample profile is from the population can be calculated by the combination of the probabilities that it belongs to each of the mixture components.

2.1 Finite Mixture Model

Finite mixture modelling is a powerful tool for density estimation and can be regarded as a flexible way to represent a probability density function (pdf). A mixture of M simple distributions (e.g. Gaussians) can be used to represent the underlying distribution of a given set of samples $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$ in $\mathbb{R}^d$. The pdf of a sample x_i can be written as

$$p(x_i|\Theta) = \sum_{j=1}^{M} \alpha_j p_j(x_i|\theta_j) \tag{1}$$

where p_j is the component density function parameterized by θ_j, α_j is the mixing proportion of each component with $\sum_{j=1}^{M} \alpha_j = 1$ and $0 \leq \alpha_j \leq 1$, and $\Theta = (\alpha_1, \alpha_2, \ldots \alpha_M, \theta_1, \theta_2, \ldots, \theta_M)$ are the parameters (mean and standard deviation for Gaussians) of the mixture. To obtain the optimal Gaussian mixture, we used the Expectation Maximization (EM) algorithm to estimate the number of sub-distributions and the parameters of each.

2.2 Improving ASM Search

In the training (modelling) stage of ASM method, the intensity profiles of a landmark P_i on all training images are modelled as a Gaussian mixture distribution characterised by the number of components M_i and parameters Θ_i, where i is the index of the landmark.

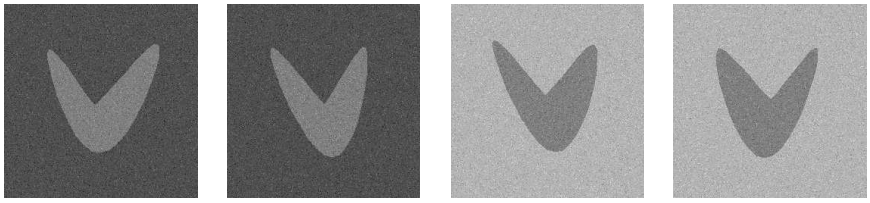
During search, we place an initial shape on the target image and, the optimal position, $\hat{P}_i$, of landmarks is the location where the local profile has maximum probability as determined by the mixture distribution.

$$\hat{P}_i \leftarrow \underset{P}{\operatorname{argmax}}\, p(x_P|\Theta_i) = \underset{P}{\operatorname{argmax}} \sum_{j=1}^{M_i} \alpha_{ij} p_j(x_P|\theta_{ij}) \tag{2}$$

where x_P is the intensity profile at position P, and P is selected along the profile across the current landmark P_i.

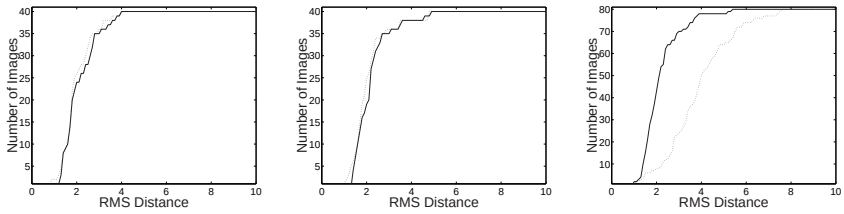
3 Experiments and Results

To evaluate the methodology, firstly a set of synthetic data, which includs two subsets, was used. Each subset includes 40 images with similar intensity for target objects, while the background intensity values in the subsets are 80 and 180 respectively. Both original and improved ASM methods were applied to individual subsets separately and then to the combination of the subsets. Subsequently, experiments were performed to segment real medical data, which consists of 10 pelvis transverse MRI sequences including 65 prostate images. The main aim of these experiments is to demonstrate the strength of the improved method when applied to images with more complex intensity distribution.



(a) (b)

Figure 1. Two example images from (a) set S_A and (b) set S_B.



(a) (b) (c)

Figure 2. The distribution of RMSD, between the ground-truth and original ASM (dotted lines), the ground-truth and improved ASM (solid lines). Here (a) set S_A, (b) set S_B and (c) set S

3.1 Evaluation Using Synthetic Images

The use of synthetic images in the testing of a method enables us to compare the results of the method with *real* ground truth. In our experiments, both original and improved ASM are tested using a same set of synthetic data, S, which includes two subsets, denoted as S_A and S_B. Each of these subsets includes 40 256×256 images. A 'V'-shape target object with shape variation is placed in the center of the images. Four key points are used to generate a nonuniform rational B-spline curve that represents the 'V'-shape boundary. To produce the shape variation, a displacement is added to each of these key points when each image is created. The displacement, (dx, dy), is randomly selected from a *Gaussian* distribution, with the standard deviation of 6 pixels. Nine points are evenly chosen on each of the four segments on the curve. Hence, 40 landmarks are used to represent the target object. The grey level intensity of the target is 128, while the background grey level for S_A and S_B are 80 and 180, respectively. Finally, we add *Gaussian* noise, with standard deviation of 16, to the images to simulate more realistic images. Example images from each subset are given in Fig. 1.

Images from S_A or S_B have little intensity variation despite the artificial noise. Set S contains all 80 images, and as such two different intensity distributions. Leave-one-image-out experiments were performed on all three sets respectively, using both original and improved ASM methods.

The Root-Mean-Square Distance (RMSD) between two shapes is used to evaluate the methods. The RMSD, from segmentation results of both ASM methods to the ground truth are compared. The distribution of RMSD is shown in Fig. 2. These results indicate equivalent performance on the image sub-sets, but a significant improvement when using the improved ASM methodology on the complete data set.

3.2 Segmentation of Prostate from MR Images

The improved ASM method is applied to a set of male pelvis transverse MRI data and the results are compared to those of original methods. The data set includes 10 pelvis transverse MRI sequences including 65 prostate images, which are obtained on a 1.5 Tesla magnet (Signa, GE Medical Systems, Milwaukee, USA) using a phased array pelvic coil, with 24×24 cm field of view, 256×512 matrix, 3 mm slice thickness and 0.5 mm inter-slice gap. All images are manually annotated by an expert radiologist and shapes are represented by landmarks. Thirty-two landmarks are used to represent the boundary of the prostate gland. Leave-one-patient-out experiments were performed on the whole data set. A number of examples of segmentation results are shown in Fig. 3.

3.3 Results

The improved ASM method performs better than the original one in all cases in our experiments. It is implied by Fig. 2 a and b that, when applied to a set of images with normal intensity variation, e.g. sets S_A, S_B, the improved method produced equivalent results to the original ASM. On the other hand, when the intensity variation are complex, such as in S, the improved method shows a significant improvement when compared with the original ASM. A clear improvement can be observed when our approach is applied to prostate segmentation, as presented in Fig. 3 and Fig. 4. When compared to the original ASM, the improved methods produced 24.6% more segmentation results with RMSD < 3.5 mm.



Figure 3. Examples of segmentation results. Original ASM (dotted lines), improved ASM (solid lines).

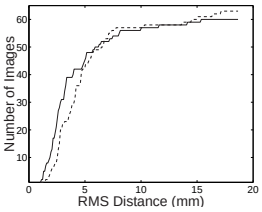


Figure 4. The distribution of RMSD, between the ground-truth and original ASM (dotted lines), the ground-truth and improved ASM (solid lines)

4 Discussion and Conclusions

As shown by the experiments, the improved method produced equivalent results to those of the original ASM when applied to images with simpler intensity distributions. Additionally, when the intensity variation are complex, the improved method shows a significant improvement when compared with the original ASM.

There is a slight deterioration in segmentation accuracy in the prostate images for both methods. A main reason for this is the less accurate boundary position and lower landmark correspondence caused by the manually annotation of prostate images, while in synthetic images the landmarks are precisely located on the real shape boundary hence the shape model and profile model have higher accuracy.

Generally, the ASM method using mixture-model framework produces faster convergence and higher robustness than the original ASM. This methodology can be applied to other choices of features used to determine the landmarks positions during ASM search, such as texture information mentioned in [5]. Since complicated intensity variation can be modelled using mixture modelling, our improved ASM method can be applied to those segmentation tasks with diverse intensity variation, such as registration in multi-modality medical imaging, and tracing objects in videos with variable object or background intensities.

Cootes et al. [2] have present their work on improving ASM by using Gaussians mixture model to represent the *shape* variation. Since we concentrate on the use of Gaussian mixture model for profile *intensity* variation, theoretically, a combination of both methods will largely improve the robustness and efficiency of ASM method. Another significant improvement to ASM was proposed by van Ginneken et al. [5], which used optimal image features and *k*-Nearest Neighbour (*k*NN) classifiers for ASM search. It is predictable that this variation can produce better results than the original ASM on the data sets we used to evaluate our method. Further work will be undertaken to make a comprehensive evaluation of ASM using mixture models for grey-level profiles compared to other variations on the basic ASM appoach.

Acknowledgements

AutoClass III in C by Cheeseman et al. [6] has greatly assisted the implementation of our algorithm and is gratefully acknowledged.

References

1. T. F. Cootes, A. Hill, C. J. Taylor et al. "The use of active shape models for locating structures in medical images." *Image and Vision Computing* **12(6)(2)**, pp. 355–366, 1994.
2. T. F. Cootes & C. J. Taylor. "A mixture model for representing shape variation." *Proceedings of British Machine Vision Conference* **1(3)**, pp. 110–119, 1997.
3. M. Rogers & J. Graham. "Robust active shape model search." *Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg New York* **2353(4)**, pp. 517–530, 2002.
4. C. J. Twining & C. J. Taylor. "Kernel principal component analysis and the construction of non-linear active shape models." *Proceedings of British Machine Vision Conference* **1**, pp. 23–32, 2001.
5. B. van Ginneken, A. F. Frangi, J. J. Staal et al. "Active shape model segmentation with optimal features." *IEEE. Transactions on Medical Imaging* **21(8)**, pp. 924–933, 2002.
6. P. Cheeseman & J. Stutz. "Bayesian classification (autoclass): Theory and results." In *Advances in Knowledge Discovery and Data Mining*, pp. 153–180. 1996.

Label propagation in MRI: potential use in diagnostic imaging

Rolf A Heckemann[a*], Joseph V Hajnal[a], Richard I Aviv[b],
Christopher J Harvey[b], Derek LG Hill[c] and Daniel Rueckert[d]

[a]Imaging Sciences Department, Imperial College; [b]Department of Imaging, Hammersmith Hospital
[c]Imaging Sciences Division, Kings College London; [d]Department of Computing, Imperial College, London

Introduction

Currently, radiological interpretation of magnetic resonance (MR) images of the brain relies mostly on subjective visual assessment. Interobserver disagreement is a problem, especially when diffuse neurodegenerative processes such as Alzheimer disease (AD) are taking place [1]. Such diseases often cause abnormalities in the size of anatomical structures, and these are particularly difficult to detect visually on sectional images. It would be desirable to support diagnostic decision making with quantitative statistical information pertaining to individual structure sizes.

In order to automatically delineate and measure individual structures, we used non-rigid registration to propagate atlas labels onto images of study subjects. We developed a novel way of presenting information thus obtained in an intuitive fashion that facilitates diagnostic decision support.

Method

We studied three-dimensional T1-weighted MR data sets from 36 subjects (7 with clinically probable or possible AD, 29 normal, age 72.5$\pm$8). A manually segmented reference MR volume [2] was used as an atlas and registered onto each study subject using a previously described approach: normalized mutual information was maximized by affine registration to correct for global mismatch, followed by high-dimensional warping to achieve congruence at the detail level [3]. The resulting transformation was then applied to the atlas label set, resulting in a new label set that corresponded to the subject's anatomy. For each labelled structure, the size was determined as a voxel count.

The AD subjects, together with seven randomly chosen normal study subjects, were used as the test subject cohort. The remaining 22 subjects were used as the reference cohort. Each member of the subject cohort was compared individually with the reference cohort, generating two per-structure comparison statistics: size rank and deviation ratio ($DR = \frac{s_{ind} - s_{average}}{\sigma}$; s_{ind}: structure size in the individual, σ: standard deviation in the reference cohort). Based on these measures, two colour overlays were generated for each test subject using pre-defined lookup tables. For the size rank overlay, a scheme was chosen where ranks 1 to 4 were shown in shades of yellow, 5 to 9 in shades of red, 10 to 13 not shown, 14 to 18 shown in shades of blue and 19 to 22 in shades of cyan. The extreme ranks were thus coloured particularly brightly. For the deviation ratio overlay, various colour schemes were tried.

The images were reviewed by a neuroradiologist and a general radiologist. Both knew the subjects' age and sex, but were blinded to the diagnosis and other clinical information. Initial impressions of AD versus normal findings from the unmodified grey scale images and diagnostic confidence on a scale from 1 to 10 were recorded. Afterwards, the size rank colour overlay was made available to be turned on and off as required by the reviewer. Diagnostic confidence was recorded again to determine the impact of the added information. An overlay was classified as *useful* if it confirmed an initially correct assessment, or if it reduced confidence in an initially incorrect assessment. If confidence in a correct assessment was reduced, or if confidence in an incorrect assessment was increased, the overlay was rated as *not useful* for the case in question. Colour overlays representing the deviation ratio were reviewed and rated by the first author.

Results

Colour overlays representing structure sizes in an individual subject as ranked with a matched reference cohort were rated *useful* in 18 of 28 cases. A *not useful* result was found in 6 of 28 cases. Table 1 shows the results by diagnosis and type of influence on diagnostic confidence (14 subjects rated by two reviewers).

In cases of AD where the assessment was correct, the reviewers' confidence was increased particularly if the overlay indicated a relatively small cortex, a small hippocampus or large lateral ventricle on at least one side. It

*Email: ixi@soundray.org

	Confidence		
	increased	unchanged	reduced
Diagnosis correct	15	3	4
Diagnosis wrong	2	1	3

Table 1. Influence of overlay information on diagnostic confidence

was noted that in all AD cases the overlay looked strongly asymmetrical (cf. figures for an example), especially in the cortices and even when the grey scale findings did not suggest a large amount of anatomical asymmetry. Asymmetry of the overlay was also seen in normal subjects.

In two of the subjects with AD, the cortex size was shown to rank high. Both of these subjects had wide sulcal spaces that were mislabelled as part of the cortex by the label propagation process. In another case of AD, massively large lateral ventricles were mislabelled as white matter. One observer rated these misregistered overlays as reducing confidence, the other observer disregarded the information pertaining to the misregistered structures and considered only the remaining, correctly registered structures.

Overlays representing the deviation ratio were rated as not providing additional benefit.

Sample images are shown in Figs. 1 (no overlay), 2 (size rank overlay) and 3 (deviation ratio overlay).

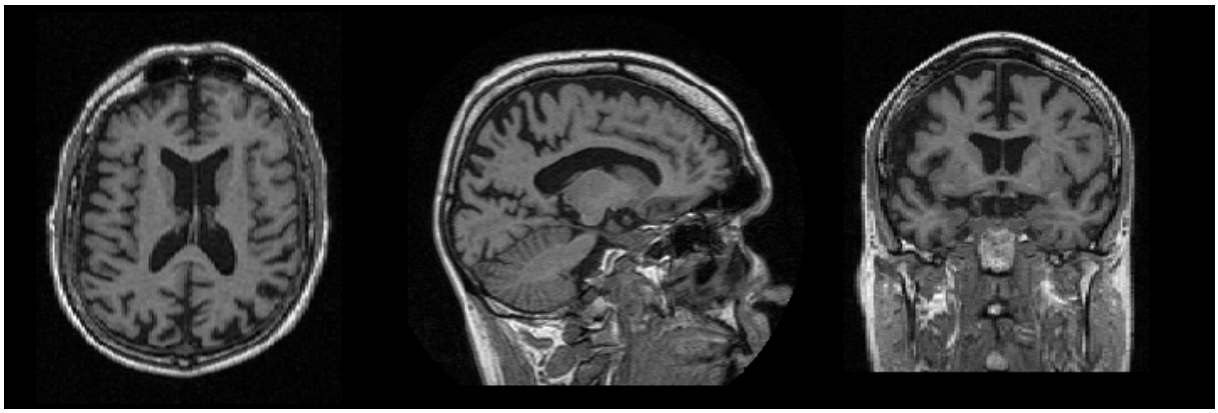


Figure 1. Unmodified MR image of the brain in Alzheimer disease. Changes are symmetric diffuse and subtle

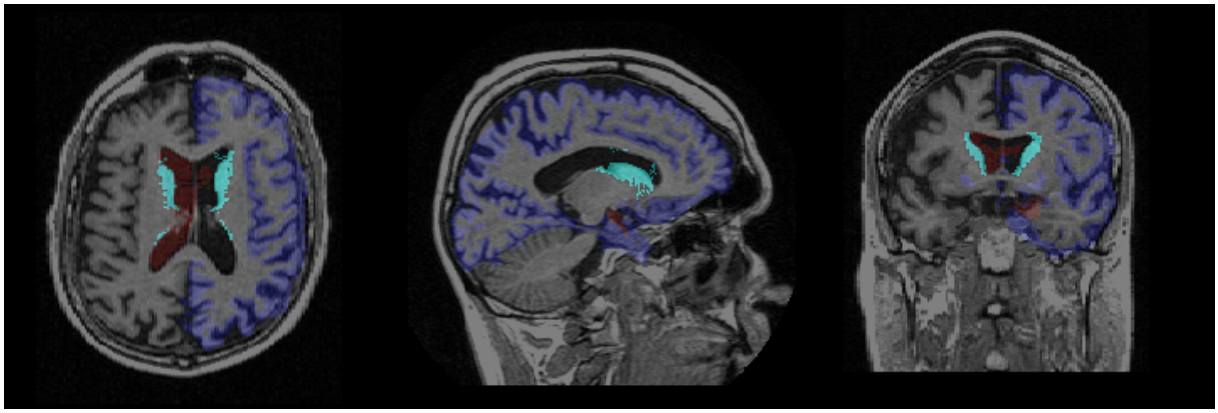


Figure 2. Same image as in Fig. 1 with size rank overlay. The overlay shows that the left cortex and bilateral ganglia are relatively small.

Discussion

Human observers are good at recognising patterns and asymmetry in images, a capability that enables them to reliably diagnose focal lesions on brain MR images. Abnormalities of structure size, however, are more difficult to spot. We used automatic anatomical segmentation based on label propagation from a manually prepared MRI brain atlas to determine structure sizes and compare them to a reference cohort.

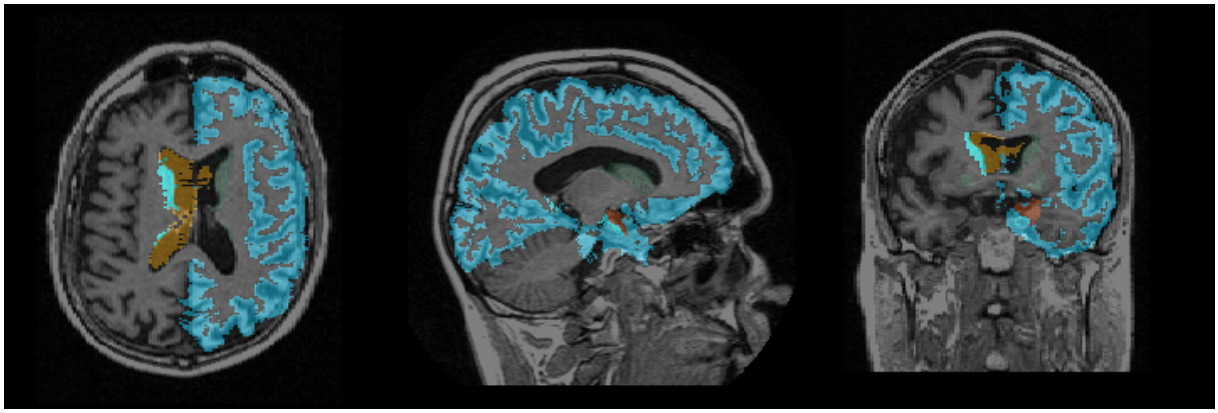


Figure 3. Same image as in Fig. 1 with deviation overlay.

We aimed to present the additional information thus gained in a fashion that integrates with image interpretation as typically performed by radiologist observers. This was achieved by providing colour overlays that can be turned on and off by the observer. Observers in this study made use of the overlays intuitively and diagnostic confidence was influenced by the additional information. The size rank information was beneficial in the majority of cases in that it either confirmed an initially correct impression, or cast doubt on an initially incorrect assessment.

Asymmetry in the overlay, especially discrepancy between the cortices, appeared to be typical for AD subjects. We intend to follow up on this observation, as it appears to be sensitive, albeit not specific for AD.

Misregistration resulted in obviously wrong rankings for some structures. The two observers dealt with the overlays in different ways: for one observer, the entire overlay became essentially unusable if obvious errors were contained, and he then noted unchanged or reduced confidence. The other observer sought to obtain as much information as possible from the overlay to support his initial assessment, in spite of these artifacts.

We expected that the other type of overlay, which represents the deviation ratio, would be even more beneficial, as the deviation ratio directly quantifies the degree of aberration from the expected size of a structure. However, on first review we found that the deviation ratio overlays did not show any advantage over the size rank overlays. We are investigating whether this contradiction can be solved by fine-tuning the lookup tables that are used to assign colours to value ranges, or whether there could be a substantial reason for size rank information to be superior.

References

1. E. van Straaten EC, P. Scheltens, D. Knol et al. "Operational definitions for the ninds-airen criteria for vascular dementia: an interobserver study." *Stroke* **34**, pp. 1907–1912, 2003.
2. D. Kennedy, P. Filipek & V. Caviness. "Anatomic segmentation and volumetric calculations in nuclear magnetic resonance imaging." *IEEE Transactions on Medical Imaging* **8**, pp. 1–7, 1989.
3. D. Rueckert, L. I. Sonoda, C. Hayes et al. "Nonrigid registration using free-form deformations: application to breast mr images." *IEEE Trans Med Imaging* **18**, pp. 712–721, 1999.

Predicting the effect of image blur on boundary shape

E Y T Ho* and C J Taylor

Imaging Science and Biomedical Engineering
Medical School, University of Manchester
Stopford Building, Oxford Road
Manchester M13 9PL, United Kingdom

Abstract. When shape boundaries are extracted from medical images, some of the details are lost due to the finite resolution of the images. We are interested in recovering super-resolution shape boundaries from sets of degraded boundaries, using an approach similar to that used to obtain super-resolution images [1]. A key step in formulating the solution is to predict the effect of image blurring directly from the boundary function. We show here that this can be achieved using a locally constant curvature approximation and demonstrate experimentally that the predictions are accurate.

1 Introduction

The idea of recovering a super-resolution image from a set of degraded images of the same scene is well established. Capel [1] describes a maximum-likelihood estimation (MLE) approach which uses a forward imaging model to predict the set of observed images, given a hypothesised super-resolution image. Because the imaging model is linear, it is straightforward to recover and estimate of the maximum-likelihood super-resolution (original) image.

We are interested in applying a similar approach to shapes. When shapes are extracted from medical images – by manual or automated segmentation – some of the details are lost due to the finite resolution of the images. A key step in developing an MLE solution to recovering super-resolution shape is to devise a forward model for the change in shape due to the combined effects of image degradation and segmentation. Ideally, we wish to predict the degraded shape boundary directly from the super-resolution boundary. To achieve this, we make the assumption that the curvature of the underlying (super-resolution) boundary is locally constant over the range of the smoothing point spread function (PSF - assumed Gaussian) involved in degrading its image.

In the remainder of the paper, we develop a theoretical relationship between super-resolution and degraded shapes, based on the locally constant curvature assumption. To test the predictions made by this model, we compare predicted change in shape with the actual change obtained by blurring a binary shape and thresholding at 0.5.

2 Theory

Given the assumption of locally constant curvature, we can model the effect of image smoothing at a point P on the shape boundary as due to convolution of the smoothing PSF with a binary disk tangent to, and with the same radius of curvature as the boundary at P. This is illustrated in **Figure 1**. To find the new position of this boundary point we need to consider the intensity at a general point, P' on the normal vector to P.

The binary disk with origin A and radius R_0 in $\Re^2$ is given by,

$$D(\overrightarrow{r}) = \begin{cases} 1 & if \parallel \overrightarrow{r} \parallel \leq R_0 \\ 0 & otherwise \end{cases} \tag{1}$$

To predict the effect of image blurring, we convolve the binary disk with a Gaussian,

$$G(\overrightarrow{r}) = \frac{1}{2\pi\sigma^2} e^{\frac{-\parallel \overrightarrow{r} \parallel^2}{2\sigma^2}} \tag{2}$$

where the Gaussian is normalised, $\int_{\Re^2} G(\overrightarrow{r}) d\overrightarrow{r} = 1$

*edward.ho@isbe.man.ac.uk

The intensity, H at point P' on the normal to the boundary at P, a distance $\overrightarrow{R}$ from the centre of the disk is given by the convolution,

$$H(R) = (D * G) = \int_{\Re^2} D(\overrightarrow{r}) G(\overrightarrow{R} - \overrightarrow{r}) d\overrightarrow{r} \tag{3}$$

Converting to polar coordinates,

$$\begin{aligned} H(R) &= \int_0^{\infty}\int_0^{2\pi} D(\overrightarrow{r}) \frac{1}{2\pi\sigma^2} e^{\frac{-\| \overrightarrow{R} - \overrightarrow{r} \|^2}{2\sigma_2}} r d\theta dr \\ &= \frac{1}{2\pi\sigma^2} \int_0^{R_0}\int_0^{2\pi} r e^{\frac{-(r^2+R^2-2rRcos\theta)}{2\sigma_2}} d\theta dr \\ &= \frac{1}{\sigma^2} e^{\frac{-R^2}{2\sigma_2}} \int_0^{R_0} r e^{\frac{-r^2}{2\sigma_2}} \left[\frac{1}{2\pi} \int_0^{2\pi} e^{\frac{2rRcos\theta}{2\sigma_2}} d\theta \right] dr \\ &= \frac{1}{\sigma^2} e^{\frac{-R^2}{2\sigma_2}} \int_0^{R_0} r e^{\frac{-r^2}{2\sigma_2}} I_0(\frac{rR}{\sigma^2}) dr \end{aligned} \tag{4}$$

where $I_0(z) = \frac{1}{2\pi}\int_0^{2\pi} e^{zcos\theta} d\theta = \sum_{k=0}^{\infty} \frac{z^{2k}}{2^{2k}(k!)^2}$ is the modified Bessel function of the first kind at zeroth-order; approximations to the function can be found in [2] and it can be evaluated readily in MATLAB.

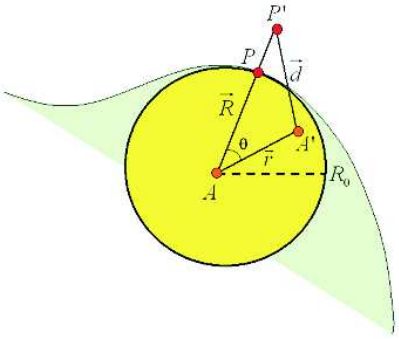


Figure 1.
A typical binary disk representation on a local curvature

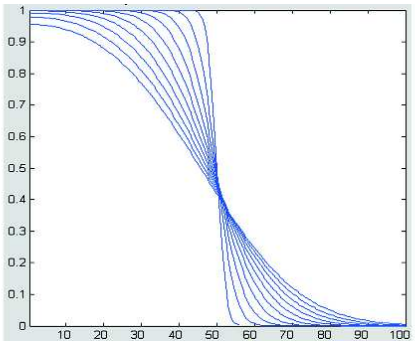


Figure 2.
Intensity versus radial distance of convolutions of a binary disk (50 pixels) with Gaussian s.d. from 2 to 20

Figure 2 shows the convolution plot of a typical binary disk (radius 50 pixels) with a set of Gaussian PSFs of standard deviation (s.d.) from 2 to 20 at intervals of 2. We can observe that at low Gaussian standard deviations, blurring only occurs at boundaries of the disks. If standard deviation of the Gaussian PSF is high enough, the entire disk is blurred and the highest intensity in the blurred version of the disk is no longer equal to the original intensity of the binary disk. The new position of the boundary at P can be found by equating $H(R)$ to 0.5 or by taking derivatives and finding the point of highest gradient. For the experiments reported below, we took the former approach because it allowed more straightforward comparison of the predicted and experimental boundaries. Note that, under this definition of the boundary, the predicted position is undefined if the ratio between σ and the local radius of curvature is too large.

If the radius of curvature is large enough (i.e. relative to the scale parameter σ, asymptotic assumptions can be made),

$$H(R, R_0, \sigma) = \frac{1}{\sigma^2} e^{\frac{-R^2}{2\sigma_2}} \int_0^{R_0} r e^{\frac{-r^2}{2\sigma_2}} I_0(\frac{rR}{\sigma^2}) dr < \frac{1}{\sigma^2} e^{\frac{-R^2}{2\sigma_2}} \frac{R_0^2}{2\sigma^2} I_0(\frac{R_0 R}{\sigma^2}) \quad (5)$$

From the asymptotic expansion of the modified Bessel function [2], $I_0(z) \approx \frac{e^z}{\sqrt{2\pi z}} \left[1 + \sum_{k=1}^{\infty} \frac{[(2k-1)!!]^2}{k!(8z)^k}\right]$

Thus for large $R_0\sigma$,

$$H(R) \approx \frac{1}{\sigma^2} e^{\frac{-R^2}{2\sigma_2}} \frac{R_0^2}{2\sigma^2} \frac{e^{\frac{R_0 R}{\sigma^2}}}{\sqrt{2\pi \frac{R_0 R}{\sigma^2}}} \left[1 + \sum_{k=1}^{\infty} \frac{[(2k-1)!!]^2}{k!(8\frac{R_0 R}{\sigma^2})^k}\right]$$

$$= \frac{R_0^2}{2\sigma^3 \sqrt{2\pi R_0 R}} e^{\frac{-R^2+2R_0 R}{2\sigma_2}} \left[1 + \frac{1}{8\frac{R_0 R}{\sigma^2}} + \frac{9}{2!(8\frac{R_0 R}{\sigma^2})^2} + \frac{9(25)}{3!(8\frac{R_0 R}{\sigma^2})^3} + ...\right] \quad (6)$$

3 Results

To test the accuracy of our prediction, we performed experiments with a synthetic binary 'limacon' shape with varying boundary curvature. We first reparameterised the shape boundary to obtain equally-spaced points around the boundary. We then found the normal vector at each point and calculated the radius of curvature. We computed movement along the normal vector at each boundary point for different values of σ using Equation (4). We also generated a binary image of the synthetic shape, smoothed it with the same values of σ and extracted the 0.5 isointensity contour. In order to reduce the effects of quantisation problems, we regularised the experimental shape boundaries using an appropriate harmonic Fourier descriptor to smooth the shape boundaries. The difference between the position predicted by Equation (4) and the experimental boundary position was measured along the normal at each point.

Figure 3 shows the original limacon shape (average diameter is 280 pixels) and its shape boundary curvature profile. We can see that at the left side of the shape boundary, the shape has positions of zero curvature (i.e. infinitely large radius of curvature).

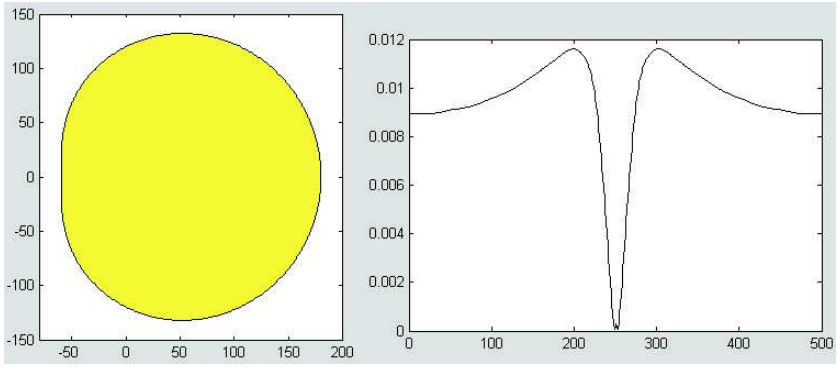


Figure 3. A binary limacon shape with continuous varying curvature profile

Figure 4 shows the experimental and predicted shape boundaries after the original image containing the binary shape is blurred using a Gaussian PSF of a standard deviation, σ of 20. We measured the root-mean square error (RMSE) of the experimental boundary with respect to the predicted and (for reference) the original boundaries. We also observed that at zero curvature positions, there is no movement for both the predicted and experimental

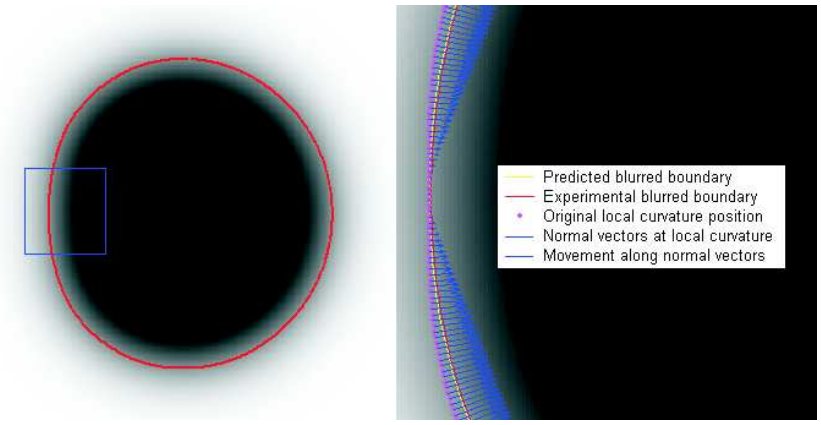


Figure 4. 0.5 isointensity contour of the experimental and predicted shape boundaries

case, as long as the Gaussian PSF has a standard deviation which is less than the zero curvature region. This is simply because the magnitudes of the normal vectors at these positions are zero.

Row 1 in **Table 1** shows the RMSE between the predicted and the experimental shape boundaries for σ ranging from 5 to 30. Row 2 shows the RMSE between the original and the experimental shape boundaries for the same range of σ. The σ were arbitrary chosen for testing purposes, as long as the highest intensity of the binary object remains higher than 0.5 after blurring. These results show that the predicted shapes are in good agreement with the experimental shapes.

σ	5	10	15	20	25	30
RMSE of Predicted and Experimental	0.5090	0.5033	0.4770	0.5113	0.4960	1.2835
RMSE of Original and Experimental	0.5411	0.8191	1.3203	2.0500	2.9933	3.5197

Table 1.
RMSE of predicted and original shape boundaries with respect to experimental shape boundaries for different values of σ

4 Discussion and Conclusion

We have shown that we can accurately predict boundary changes due to image blurring under certain assumptions. Principally, we assume that the underlying (super-resolution) shape boundary has locally constant curvature over the range of the Gaussian PSF used for image smoothing. This assumption is fundamental, but is particularly important in our experimental regime where the 0.5 isointensity contour becomes undefined if the the assumption does not hold. We also assume that the size of the image is large enough that we do not need to worry about boundary effects. The assumption of locally constant curvature is particularly problematic in regions such as cusps, where the curvature changes rapidly. This might be tackled by using a higher order approximation to the local shape of the curve [3].

Acknowledgements

The authors would like to thank Dr J. Graham for some constructive comments.

References

1. D.P. Capel. *Image Mosaicing and Super-resolution.* PhD thesis, University of Oxford, 2001.
2. M. Abramowits and I.A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, New York, 1965.
3. F. Bergholm. *Edge Focusing.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.9, No.6, Nov. 1987.

Orientation Coherence Optimisation in Tensor Image Registration

KM Curran and DC Alexander *

Computer Science Department , University College London, Gower Street, London WC1E 6BT, UK.

Abstract. Registration of diffusion-tensor magnetic-resonance (DT-MR) images is distinguished from single component intensity image matching by the orientational information the images contain, which is affected by the registration transformation. We describe novel DT-MR registration techniques that have been devised to exploit the orientation coherence of diffusion tensors and to determine if diffusion-tensor-orientation matching improves the accuracy of registration over that from registration algorithms that ignore orientation. We show comparative results from inter-subject human brain data. To compare the potential of orientation-based methods fairly with standard methods we use combined simulated annealing and gradient descent to get closer to the global minimum of the registration objective function.

1 Introduction

Diffusion-tensor MRI [1] measures the diffusion tensor of water molecules in each image voxel. White matter fibres hinder water mobility more across the fibre than along it so that the principal axis of the diffusion tensor is much larger than the other two axes, and it coincides with the direction of the fibers. This reveals structural information about anatomical regions that otherwise appear homogenous on conventional MR images. Applications of DT-MRI lie mostly in the analysis of pathologies of white matter in the brain, for example, multiple sclerosis [2] and schizophrenia [3]. Tractography algorithms use fibre-orientation estimates from DT-MRI to follow fibre trajectories through image volumes. Tractography helps neurosurgical planning [4] and clinical studies into differences in connectivity between different population groups [5]. Such studies require spatial normalisation to remove variability in fibre tract position, shape and thickness. Here we investigate registration of diffusion-tensor magnetic-resonance (DT-MR) images. In particular, we test the hypothesis that diffusion-tensor orientation matching improves the accuracy of registration over registration algorithms that ignore orientation.

Image-based registration of scalar images estimates a transformation that minimises the difference between the target and source images. Alexander and Gee [6] extend an elastic matching algorithm, originally designed for scalar imagery, to work with DT-MR images. They describe a number of similarity measures, derived from the full tensor, that are sensitive to orientational information. Other similarity measures have been proposed by Ruiz-Alzola [7] and Guimond [8].

Accurate registration of diffusion tensor images also requires tensor reorientation to keep their orientations consistent with the image structure [9]. If the transformation is rigid, we can reorient the tensors directly: if R is the rotation matrix of the image transformation, each $\mathbf{D}$, becomes $\mathbf{D}'$, where $\mathbf{D}' = R\mathbf{D}R^T$. For higher order transformations, we must determine the rotational effect that the transformation has on the tissue microstructure and use that rotation to reorient the diffusion tensor. A number of reorientation strategies have been proposed, including finite strain (FS) and preservation of principal direction (PPD), which try to find rigid transformations that reflect the local reorientation of the image from the transformation [9]. Any non-singular affine transformation F can be decomposed into a rigid component R and a deformation component U [11] so that $F = UR$. The FS strategy [9] uses the rigid component, R, of the local affine approximation, F, to the transformation to reorient $\mathbf{D}$. We can compute R from F using $R = (FF^T)^{-1/2}F$. If the transformation is affine, the rotation R needs to be computed only once and is constant over the entire image. A drawback of the FS strategy is that the deformation component U of the affine transformation is discarded and does not contribute to the estimated reorientation [9]. The PPD strategy directly examines the effects of the transformation on the eigenvectors of the DT at each point. The PPD algorithm finds the rotation that maps $\mathbf{e_1}$, the principal eigenvector of $\mathbf{D}$, exactly to $F\mathbf{e_1}$ and maps $\mathbf{e_2}$, the second eigenvector, closest to $F\mathbf{e_2}$. The strategy assumes that the directionality of the tissue structure corresponds to the direction of the eigenvectors of the DT [12]. With PPD, the DT reorientation is not constant over the image for general affine transformations, unlike the FS strategy and a separate R must be computed at each voxel. Xu and colleagues [13] refine the PPD approach by estimating the fibre direction more reliably from a neighbourhood of voxels.

In earlier work [14], we adjust the tensor orientations iteratively during registration according to the transformation using the FS or PPD method. We show that orientation matching improves registration results using direct

*K.Curran@cs.ucl.ac.uk and D.Alexander@cs.ucl.ac.uk

optimisation for intra-subject simulations. Park and colleagues [15] also use orientation matching to drive the registration and they found that full-tensor registration with PPD reorientation showed the best performance in effectively normalising the tract morphology and tensor orientation. Their study introduces evaluation measures based on alignment of extracted fiber tracts from the registered data. They acknowledge that errors in fiber tracking affect the evaluation of the registration but suggest that such errors are evenly distributed for all comparison procedures. Here we examine the directional coherence of the transformed-source and target images in two hand-defined regions. Experimentation with the direct optimisation scheme used in [14] within inter-subject brain registration tasks revealed difficulties with local minima. An additional contribution of this work is the highlighting of the local minimum problem in orientation matching and the need to use global optimisation techniques, despite the dramatic increase in computation times, to overcome these difficulties. Finding optimal solutions is particularly important here, since the aim of the work is to determine which approach fundamentally provides the best image match. Spurious optima at local minima will not demonstrate the full power of a particular matching technique.

In section 2 we outline the DT-MR registration algorithm. Section 3 presents the experiments and results. Finally, we conclude in section 4 and discuss directions for future work.

2 Method

In this paper, we limit investigation to affine transformations, although the methods we describe can be extended easily to other transformation groups. We use a standard decomposition of the affine transformation, which decouples the parameters, splitting the transformation into 3 rotation, 3 translation, 3 scale and 3 skew parameters. We sum the voxelwise similarity over the overlapped foreground regions of the transformed source and target images and normalise by the size of the overlap. We consider only an eight of the voxels ($32 \times 32 \times 42$) in order to reduce computation times. We can use any reorientation strategy to compute the transformed source image. For tensor images, we can choose the similarity to be any of the indexes described in [9]. We use the tensor difference Δ, which is the sum of square differences between the nine corresponding elements of the two DT matrices. For two DTs, $\mathbf{D}_1$ and $\mathbf{D}_2$, $\Delta(\mathbf{D}_1, \mathbf{D}_2) = \sqrt{(\mathbf{D}_1 - \mathbf{D}_2) : (\mathbf{D}_1 - \mathbf{D}_2)}$, where the tensor scalar product $\mathbf{D}_1 : \mathbf{D}_2 = \mathrm{Tr}(\mathbf{D}_1\mathbf{D}_2)$. The tensor difference is sensitive to differences in size, shape and orientation of the two tensors [9] and has proved effective for DT image matching in previous work [6, 9].

Optimisation methods seek the global minimum of an objective function. However, a common problem with optimisation techniques is that they can converge to sub-optimal solutions at local minima. In an attempt to find the global minimum in a reasonable amount of time, many registration methods rely on multi-resolution approaches in the hope that local optimisation will find the global minimum. Jenkinson et al [16] present a global optimisation for volumetric registration of brain images. This method combines a fast local optimisation, Powell's method [17], with an initial search phase, tuned to be computationally feasible. In a similar way, to avoid being trapped in local minima and thus sub-optimal solutions, we combine a fast local optimisation, Powell's method, with a global optimisation technique, Simulated Annealing [17].

Simulated Annealing [17] is a global optimisation technique that has been reported to perform very well in the presence of a very high number of variables [18]. It is based on random evaluations of the objective function, in such a way that transitions out of a local minima are possible. It does not guarantee to find the global minimum, but if the function has many near-optimal solutions, it should find one. Here, we use simulated annealing to optimise the starting point for Powell's method. We call the combined method *gradient annealing*. We omit full details of the method, but we choose settings that give good results in manageable computation times.

3 Experiments and Results

We registered six DT-MR brain images ($128 \times 128 \times 42$ voxels) to a seventh, template image, using the algorithm described in Section 2. Using the gradient annealing optimisation scheme, we compute the affine registration using no reorientation, FS and PPD reorientation. The hypothesis is that registration using FS and PPD should be better than using no reorientation, since these methods allow orientations to be matched properly. Furthermore, PPD should produce better results than FS, since it reorients tensors more accurately.

Table 1 compares the final summed Δ after optimisation using a single run of Powell's method initialised at the identity transformation with that from the full gradient annealing optimisation. The results in Table 1 show that the direct method clearly does not find the global minimum. The gradient annealing algorithm is not guaranteed

Table 1. The summed tensor difference Δ found for the Powell initialised with an identity transformation and gradient annealing DT-MR brain registrations for each of the three affine reorientation strategies. The lower scores indicate greater similarity.

Registration Scheme	**Powell**			**Gradient Annealing**		
Reorientation Strategy	**None**	**FS**	**PPD**	**None**	**FS**	**PPD**
1	475.63	452.26	477.87	451.82	448.13	448.63
2	424.26	425.02	443.29	422.17	421.61	424.63
3	425.18	424.95	426.93	424.92	423.51	424.16
4	430.45	455.49	457.68	421.97	423.82	423.43
5	422.36	418.73	419.21	413.78	412.76	414.81
6	452.48	456.37	457.37	430.35	440.61	430.59

to have found the global minimum either, but the minima it finds are considerably closer to the global minimum. The global optimisation reduces the final objective function value consistently and by as much as 6 percent.

The final similarity measure is not a reliable indicator of registration quality. To assess the registration quality, we examine the directional coherence of the transformed-source and target images in anatomic regions outlined by hand in the template image. The selected white-matter regions are the cortico-spinal tract (CST) and the corpus callosum (CC), where the diffusion tensor is consistently anisotropic and prolate. The hand segmentation is conservative to ensure that only the chosen tissue type is included in each region of interest (ROI). The better the images are registered, the smaller the errors in tensor alignment.

In each ROI, we compute the average angular separation [9] of the principal DT eigenvectors over the ROI. The average angular separation is

$$E = \left(\sum \sqrt{\nu_1 \nu_2}\right)^{-1} \left(\sum \left(\sqrt{\nu_1 \nu_2} \times \cos^{-1} |\mathbf{e_1} \cdot \mathbf{e_2}|\right)\right) \tag{1}$$

where ν_1 and ν_2 are the anisotropies of the two diffusion tensors, $\mathbf{e_1}$ and $\mathbf{e_2}$ are their principal eigenvectors and the sums in (1) are over all the voxels in the ROI. We use a measure of prolateness [19] for ν in (1) defined by:

$$\nu = \left(\mathrm{Tr}(\mathbf{D}^2)\right)^{-\frac{1}{2}} (\lambda_1 - \lambda_2) \tag{2}$$

where $\lambda_1 > \lambda_2 > \lambda_3$ are the eigenvalues of the diffusion tensor **D**.

Table 2. The orientation statistics for the DT-MR brain registrations using no reorientation (NONE), finite strain (FS) and preservation of principal direction (PPD) reorientation.

Tissue Type	**Corpus Callosum**			**Cortico Spinal Tract**		
Reorientation Strategy	**None**	**FS**	**PPD**	**None**	**FS**	**PPD**
1	0.06	0.07	0.09	0.16	0.11	0.09
2	0.24	0.24	0.27	0.13	0.13	0.13
3	0.15	0.13	0.16	0.12	0.11	0.13
4	0.20	0.14	0.17	0.18	0.21	0.10
5	0.11	0.13	0.22	0.19	0.12	0.10
6	0.16	0.16	0.14	0.13	0.10	0.09

Table 2 shows the average angular separation of the principal eigenvectors for the six subjects for each of the reorientation strategies and in each ROI. Low numbers indicate good coherence and high numbers indicate poor alignment of the principal direction between tensors in the source and target images. The PPD method produces smaller angular separations than the FS method and both FS and PPD produce smaller angular separations than the no reorientation method for the CST region for subjects 1, 5 and 6, which is the expected pattern. The angular separations are equal for all three reorientation strategies for the CST region for subject 2. However, for subjects 3 and 4 the PPD method does not align the principal directions better than FS or no reorientation in the CST region, although for subject 3, FS performs better than the no reorientation method. However, for the CC region, the pattern is not the expected one and PPD and FS appear consistently worse than no reorientation. Closer

examination reveals that the quality of registration obtained from the affine transformation is not as good for the CC region compared with the CST region for all subjects. The poor alignment in this region gives rise to the higher than expected angular separations.

4 Conclusion

We wish to investigate the potential improvements gained from orientation matching but are not proposing a working registration algorithm because while the scheme improves results over single-shot direct optimisation, it also dramatically increases computation times. The results presented here are promising, suggesting that orientation matching improves registration to some extent. We may see a more dramatic effect for higher-order transformation models, which should improve the alignment between subjects. Future work will investigate how to incorporate orientation matching into faster algorithms reliably.

References

1. P. Basser, J. Matiello & D. L. Bihan. "MR diffusion tensor spectroscopy and imaging." *Biophysical Journal* **66**, pp. 259–267, 1994.
2. M. Cercignani, M. Inglese, E. Pagani et al. "Mean diffusivity and fractional anisotropy histograms of patients with multiple sclerosis." *American Journal of Neuroradiology* **22(5)**, pp. 952–958, 2001.
3. I. Agartz, J. Anderson & S. Skare. "Abnormal brain white matter in schizophrenia: a dt-mri imaging study." *Neuroreport* **12(10)**, pp. 2251–2254, 2001.
4. C. Poupon, C. Clark & V. Froulin. "Regularisation of diffusion-based direction maps for the tracking of brain white matter fasciles." *NeuroImage* **12**, pp. 184–195, 2000.
5. O. Ciccarelli, A. Toosy, G. Parker et al. "Diffusion tractography based group mapping of major white-matter pathways in the human brain." *Annals of Neurology. In Press* 2003.
6. D. Alexander & J. Gee. "Elastic matching of diffusion tensor images." *Compuer vision and image understanding - Special issue on analysis of volumetric images* **7(2)**, pp. 233–250, 2000.
7. J. Ruiz-Alzola, C. Westin, S. Warfield et al. "Nonrigid registration of 3d scalar vector and tensor medical data." *Medical Image Computing and Computer-Assisted Intervention-MICCAI'00* 2000.
8. A. Guimond, C. Guttmann, S. Warfield et al. "Deformable registration of DT-MRI data based on transformation invariant tensor characteristics." *In Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI'02, Conference held in Washington (DC), USA, July 7-10 2002)* 2002.
9. D. Alexander, C. Pierpaoli, P. Basser et al. "Spatial transformations of diffusion tensor images." *IEEE Transactions on Medical Imaging* **20(11)**, pp. 1131–1139, 2001.
10. J. Ruiz-Alzola, C. Westin, S. Warfield et al. "Nonrigid registration of 3D tensor medical data." *Medical Image Analysis* 2002.
11. L. Malvern. *Introduction to the mechanics of a continuous medium.* Prentice Hall, Englewood Cliffs, N.J., 1969.
12. P. Basser & C. Pierpaoli. "Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI." *Journal of magnetic resonance* **Series B 111**, pp. 209–219, 1996.
13. D. Xu, S. Mori, M. Solaiyappan et al. "A framework for callosal fiber distribution analysis." *Neuroimage* **14**, pp. 1361–1369, December 2001.
14. K. Curran & D. Alexander. "Diffusion tensor orientation matching for image registration." *Proc. SPIE Vol. 5032, Medical Imaging 2003: Image Processing; Milan Sonka, J. Michael Fitzpatrick; Eds.* pp. 149–156, 2003.
15. H. J. Park, M. Kubicki, M. E. Shenton et al. "Spatial normalization of diffusion tensor MRI using multiple channels." *Neuroimage* **20(4)**, pp. 1995–2009, 2003.
16. M. Jenkinson & S. Smith. "A global optimisation method for robust affine registration of brain images." *Medical Image Analysis* **5(2)**, pp. 143–156, 2001.
17. W. Press, S. Teukolsky, W. Vetterling et al. *Numerical recipes in C. (The art of scientific computing).* Cambridge University Press, 2nd edition, 1992.
18. L. Ingber. "Simulated annealing: Practice versus theory." *Journal of Mathematical and Computational Modelling* **18**, pp. 29–57, 1993.
19. C. Westin, S. Maier, H. Mamata et al. "Processing and visualization for diffusion tensor MRI." *Medical Image Analysis* **6**, pp. 93–108, 2002.

Cellular Algorithms for Medical Image Processing

Luis Ibáñez

Kitware, Inc., Clifton Park NY 12065, USA

1 Introduction

This paper presents a discussion on the application of biologically-inspired algorithms to the processing of medical images. In particular, the concept of *Cellular Algorithms* is described and its relationship to other algorithms commonly used in image processing is discussed. Cellular Algorithms are based on emulating the basic characteristics of the behaviour exhibited by biological cellular aggregates. Given that cellular aggregates are the constitutive elements of biological tissues, and thus are the bases of anatomical organs and systems, their behavioural mechanisms are *naturally* suited for representing biological shapes. Therefore, algorithms emulating the behaviour of cellular aggregates offer a powerful tool for performing segmentation of anatomical structures from medical images. This paper presents the software framework under which families of Cellular Algorithms can be implemented in the Insight Toolkit (ITK) and can be made available to the medical image analysis community.

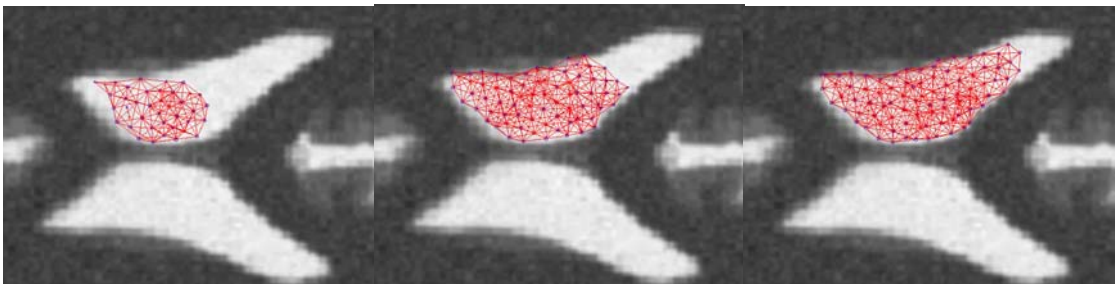


Figure 1. Illustration of an image segmentation method based on an artificial Bacterial Colony. The mesh is in this case the output of the segmentation process. The nodes represent Bacteria. Snapshots were taken at three different times of the execution

2 Background

Cellular Algorithms are not new at all; they have been in use since the very early years of image processing. In fact, most of the techniques currently used in image processing are specializations of Cellular Algorithms. This includes algorithms such as Mathematical Morphology, Neighbourhood Filtering, Convolution Filters, Variable Conductance Diffusion Filters, Markov Random Fields, PDE Methods, FEM methods, Level Sets, Region Growing, Deformable Models, and Watersheds. Unfortunately, the theoretical algorithmic aspects behind these commonly used methods have been usually neglected or disregarded, partially for the desire of presenting them as *Mathematical* approaches rather than the *Heuristics* that they actually are.

The fundamental characteristics of Cellular Algorithms are

The Field: A set of similar elements distributed in space.
The Element State: Every element in the field has a state that is taken from a set of possible states.
The Rules: The field is iteratively updated based on a rule that is applied to every element.
The Rule Scope: The update rule takes into account the state of other elements in previous iterations.

At this point the reader probably recognizes the signature of Cellular Automata. John Von Neumann under the advice of Stanislaw Ulam, was one of the first people to consider the model of Cellular Automata in his seminal work on Self-Reproducing Automata [1]. John Holland started applying cellular automata to problems of adaptation and optimisation [6]. Strictly speaking, Cellular Automata are equivalent to Cellular Algorithms, however, technological limitations of the early days of computing led to implementing Cellular Automata in very restrictive contexts. Subsequent evolution of these algorithms lacked of historical perspective, and resulted in the continued imposition of such restrictions despite the fact that our current computational capabilities do not

Corresponding author: luis.ibanez@kitware.com, KITWARE Inc, 28 Corporate Rd, Clifton Park, NY 12065

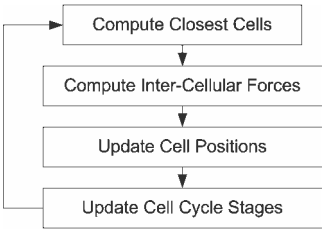
require them anymore. Among those restrictions, the set of elements has traditionally been represented as the nodes of a rectilinear grid, the position of the elements does not change as a consequence of the application of the rule, and nodes in the grid are not inserted or removed. These implementation restrictions have been associated so much to the definition of Cellular Automata that we have preferred to use here the term Cellular Algorithms in order to return to the original concept of emulating the behaviour of a group of biological cells. The contribution of this paper is therefore to put current image processing algorithms into perspective and to explore the implementation of new ones that take advantage of cellular behaviours underexploited so far.

3 Algorithms and Computational Capacity

A large number of image processing methods currently in use for medical applications are based on the paradigm of visiting all the pixels on the image and updating their values according to a particular rule. The rule typically involves the evaluation of the pixel neighbours. In some cases, the processing requires several input images and therefore the computation rule requires the evaluation of pixels in various input images. This generic paradigm is the basis of simple image processing such as mean and median filters, as well as more complex algorithms such as those of Mathematical Morphology, and those of Region Growing methods for segmentation. The paradigm above also applies to methods solving Partial Differential Equations (PDE) iteratively, which includes advanced algorithms such as the Demons deformable registration methods and all the implementations of Level Sets for image segmentation, image enhancement and motion detection. Statistical methods such as Markov Random Fields also belong to this general group with the interesting variation of introducing stochastic rules. Other methods are not attached to the image grid, but still follow the generic approach of having a set of basic elements that are iteratively updated by the application of a rule in which the state of other elements is taken into account. This is the case of FEM-based registration algorithms and Deformable Models such as Snakes used for image segmentation. From the Algorithmic Theory point of view, all the methods cited above are a very narrow subset of the computational entities that can be modelled with Cellular Automata (CA) [1, 2, 3, 4]. In fact, all the image processing implementations cited above ***are*** Cellular Automata. A significant advance in image processing may be achieved by using more generic implementations of Cellular Automata, in particular those based on non-regular grids and dynamic grids where insertion and deletion of elements is possible. These more generic algorithms are inspired in the behaviour of biological cells and are referred here as *Cellular Algorithms* [5].

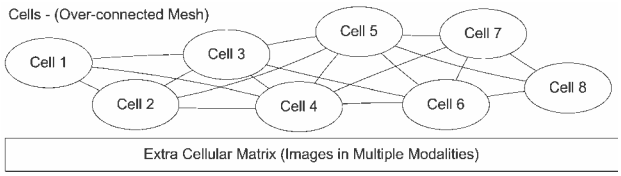
The algorithmic capabilities of Cellular Automata have been studied in detail for several decades. Let's just remark that very simple Cellular Automata have been proved to be able to perform as "*Universal Computers*". This means that they can be used for implementing *any* algorithm. However, little progress has been made on the application of their generic formulation. This is mainly due to the difficulty for anticipating the behaviour that results from the definition of a particular updating rule. Small rule changes that may appear insignificant, often times result in dramatic transformations on the behaviour of the field over time. The selection of the update rules is the most critical element on the definition of Cellular Algorithms.

3 Bacterial Colonies for Segmentation



A generic framework for supporting Cellular Algorithms is being implemented in the Insight Toolkit (ITK). Out of this generic framework, a simple algorithm for image segmentation is illustrated here by simulating the basic rules of a Bacterial Colony. In this particular case all cells are equivalent and the image itself is presented as the substrate in which the cells are living. Cells are programmed to reproduce only when the intensities of the substrate are in a certain range. The algorithm is started by placing an initial bacterium in an image position and then letting it reproduce under the control of its own cell cycle rules. After cellular division the cells grow and push each other expanding the cellular aggregate as a result. In this simple example, cells are not actively migrating nor exhibiting any tropic behaviour. However, the addition of those behaviours is certainly a possibility for implementing more powerful variations of segmentation algorithms. The cells are represented in a Mesh where the nodes contain information about the cell state and the edges represent the neighbourhood connections. For the sake of efficiency, this mesh is implemented as an over-connected mesh, which means that a cell has connection to the immediate closest cells and also to the next set of neighbour cells. Since the cells positions change as a result of the inter-cellular forces as well as and the division of cells undergoing mitosis, it is necessary to periodically update the list of neighbours of every cell. This is done here every ten iterations, under the observation that a certain number of iterations of force computation are required for bringing the system to equilibrium after a cell division. Once the new neighbourhood relationships are updated, new inter-cellular

forces are computed, cells positions are updated, and the state of every cell in its own cycle is updated. As a result of updating the cellular cycle, some cells will undergo division; some other may undergo apoptosis, also known as *programmed cell death*. One iteration of the algorithm is illustrated in the figure above.

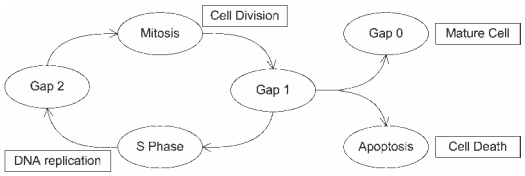


The use of an over connected mesh facilitates to recomputed the set of immediate neighbours participating in the force interaction computations. The assumption is that under steady conditions of the mesh evolution, only cells already in the second range of neighbours will eventually become closer an enter the first range of closest neighbours. Thanks to this, it is not necessary to explore the entire mesh in order to refresh the nearest neighbours' connections, making that a traditional N^2 process becomes an MxN process, where N is the number of cells in the colony and M is the number of neighbours expected in the first and second closest range. This factor is dependent on the dimension, for example, M will be about 18 in 2D, and close to 40 in 3D. The reason why the value can only be estimated is that the cells will arrange themselves in irregular grids and therefore only a mean number of neighbours can be considered here.

4 Fundamental Cellular Mechanisms

This section describes the main mechanisms responsible for controlling the behaviour of biological cellular aggregates. Some of those mechanisms have already been implemented in the generic framework of cellular aggregates in ITK, some others are only planned as future work.

4.1 Cell Cycle



The cell cycle is one of the fundamental mechanisms for controlling the behaviour of cellular aggregates. The cell cycle is the process by which cells reproduce through division, stabilize or die. The regulation of the cell cycle allows controlling the local density of cells by triggering tissue growth or by removing unnecessary cells. The loss of control over the cell cycle results in disorganized growth such as malign tumours [8]. The simple segmentation method illustrated in Figure 1 uses the intensity of the image as one of the criteria for controlling the cell cycle. Cells are only replicated in the regions of the image with bright intensities. This simple rule results in "*Region Growing*" behaviour for the cellular aggregate. The advantage of using cells here is that a large number of extra elements can be added to the decision checkpoint controlling whether the cells should replicate or not. Those extra elements may help to prevent local leaks which are the major drawback of regions growing algorithms, including those already available in ITK.

4.2 Cell-Signalling

About 40% of the proteins in unicellular organisms are dedicated to information processing and transmission tasks. That is, they do not have any structural or metabolic role in the cell life. Their purpose is to convey information from one site of the cell to another, from one cell to another cell, and between the cell and the extra cellular matrix. The molecular mechanisms of cellular communication usually involve proteins that act as signals and receptors. In many cases, those proteins are simultaneously signals and receptors and are capable of performing basic information processing tasks such as AND and OR gates [8]. In the context of software engineering, this is equivalent to an "event-driven" architecture, where cellular signals are equivalent to *Events* and cellular receptors are equivalent to *Observers* or *Listeners*. Although Events and Observers are already implemented in ITK for the control of the data-pipeline, their implementation is not appropriate for being used in the context of a large population of cells. A new mechanism for rapid processing of events from thousands of cells is expected to be implemented in ITK. One of the typical mechanism by which bacterial colonies achieve group behaviour is through "*voting"* or *"quorum sensing"* schemes. This is done by making all the bacteria emit signals that have a limited range and/or a limited lifespan. The cells have receptors for the same signal and can measure the concentration of the signalling molecule in the local neighbourhood. When the concentration of the signal surpasses a threshold, the cell interprets that as reaching a quorum amongst the neighbours and therefore proceeds to trigger a particular behaviour. This mechanism is used by fluorescent bacteria in order to pulse at night in the ocean by synchronizing large scale populations of cells [7].

4.3 Cell-Cell Interactions

Mechanical interactions between cells involve the application of pressure forces resulting from pushing cells against other, as well as tension forces resulting from adhesion between cells, and between cells and the extra cellular matrix [9,10]. These mechanical interactions are being currently simulated in ITK in the context of a highly viscous medium, where inertia has a minimum effect and cell displacements occur at low speeds. Interaction between a cell and its neighbours involve both information transmission and mechanical interactions. Information transmission is performed with signals and receptors in the cell surface. Mechanical interactions are achieved through the synthesis of proteins called *cadherins*. The cell produces these proteins and exports them to the cellular membrane from where they can attach to complementary molecules on the surface of neighbouring cells. A large variety of cadherins and complementary proteins allow the cells to selectively attach to certain types of cells and not to others. This mechanism facilitates the creation of epithelial cells and endothelial cells. Epithelial cells tend to form on the outside of organs and arrange themselves on sheets that recover the organ. Endothelial cells on the other hand stay in the inside of organs. From the algorithmic point of view, the strength of the cadherins interactions joining the epithelial cells among themselves is responsible from controlling the curvature of the cellular aggregate borders. In this way cellular aggregates can regulate their curvature in the same way that deformable models use internal forces, and Level Set methods use curvature terms in the computation of their speed images. The advantage of cell-cell interaction is that the rules controlling cadherins expression allow for more complex behaviours. For example, it is possible to set a higher curvature restriction in one side of the organ that the other, depending on what organs are found across the epithelium [8,9,10]. Currently only pushing forces are implemented in the ITK framework. Future work will introduce the emulation of cadherins.

4.4 Cell-Matrix Interactions

The cellular matrix is a substrate in which the cells live. The material of the cellular matrix is usually produced and maintained by the cells themselves, which results in a symbiotic relationship. The interactions between the cells and the matrix include information transmission via signals and receptors, metabolic dependencies due to transport of materials required for maintaining the infrastructure of both intra and extra cellular spaces, and mechanical interactions driven by proteins that attach the cytoskeleton and the cell membrane to fibres in the extra cellular matrix [8,9,10]. A natural way of using cellular algorithms for image processing is to associate the image to one of the chemical compounds of the cellular matrix. Cells in the aggregate will then react to the different concentrations of the substrate, in other words, the result of applying the update rules will depend on the intensity of the input image or images in the neighbourhood of the cell location.

Source Code

The source code of the current implementation of Cellular Algorithms in ITK is available at http://www.itk.org as part of *InsightApplications* package. They are located in the *Morphogenesis* subdirectory.

References

1. J. von Neumann. *Theory of Self-Reproducing Automata.* University of Illinois Press, 1966.
2. A. Turing. "The chemical basis of morphogenesis", in *Morphogenesis*, P.T. Saunders, Elsevier Science, 1992.
3. A. Rosenfeld. *Picture Languages – Formal Models for Picture Recognition.* Academic Press, 1979.
4. S. Wolfram. *A New Kind of Science.* Wolfram Media, Bellingham, 2002.
5. R. Nagpal. "Programmable Self-Assembly using Biologically-Inspired Multiagent Control", *Autonomous Agents and Multiagent Systems Conference* (AAMAS), July 2002.
6. J. H. Holland. "Studies of the Spontaneous Emergence of Self-Replicating Systems Using Cellular Automata and Formal Grammars", in Lindenmayer, A. and Rozenberg, G. (eds.) Automata, Languages, Development (New York, NY: North-Holland Publishing), pp. 385-404.
7. More MI, Finger LD, Stryker JL, Fuqua C, Eberhard A, Winans SC. “Enzymatic synthesis of a quorum-sensing autoinducer through use of defined substrates.”, Science. 1996 Jun 14;272(5268):1655-8.
8. H. F. Lodish, A. Berk, P. Matsudaira, C. A. Kaiser, M. Krieger, M. P. Scott, S. L. Zipursky, J. Darnell, “Molecular Cell Biology”, Fourth Edition, W.H. Freeman, 1998.
9. G.M. Odell, G. Oster, P. Alberch, B. Burnside, "The Mechanical Basis of Morphogenesis", Developmental Biology, Vol. 85, 1981, pp. 446-462.
10. S. Gilbert, “Developmental Biology”, seventh edition, Sinauer Associates, 2003.

Flaws in Existing Spatio-Temporal and Temporo-Spatial Realignment of FMRI Data

Peter R Bannister[a], J Michael Brady[a] and Mark Jenkinson[b*]

[a]Robotics Research Group, Department of Engineering Science,
University of Oxford, Parks Road, Oxford OX1 3PJ, UK.
[b]Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB),
John Radcliffe Hospital, Oxford OX3 9DU, UK.

1 Introduction

It is widely recognised that the separate application of rigid-body motion correction and temporal interpolation (slice-timing) do not accurately account for the slice-based acquisition of FMRI data. We demonstrate that the use of volumetric motion correction and slice-timing is not only inaccurate, but that in many situations it can lead to a degradation of the images.

Furthermore, knowing that the separation of motion correction and slice-timing correction is an approximation, there is no clear consensus as to what order these two steps should be applied in. We show that over a very similar set of motions, either combination can lead to a range of different corrections.

2 Existing Approaches

Echo Planar Images are typically acquired in a slice-sequential manner, that is, all the voxels falling within a single slice of scanner space are captured before the voxels in the next slice. While acquiring more than one slice at a time is possible in some MRI applications, it is currently impractical for FMRI because the resulting signal-to-noise ratio is too low to be usable.

Single-shot EPI (as opposed to multi-shot, interleaved or segmented EPI) is the EPI sequence that is typically used in FMRI. The 'single shot' refers to the fact that each slice is reconstructed from a single RF excitation. Although it is possible to acquire several slices with single shot — one shot for each slice — the FMRI data-sets acquired to date are nearly always single-shot multi-slice EPI.

Given that SNR considerations limit FMRI to stacked-slice acquisitions in practice, the problem of temporal offsets within a volume due to the successive acquisition of slices remains a significant confound to motion correction. Previous attempts to correct FMRI data for artefacts introduced during acquisition have considered spatial realignment and slice-timing correction as two distinct and separate stages in the processing chain as shown in Figure 1.

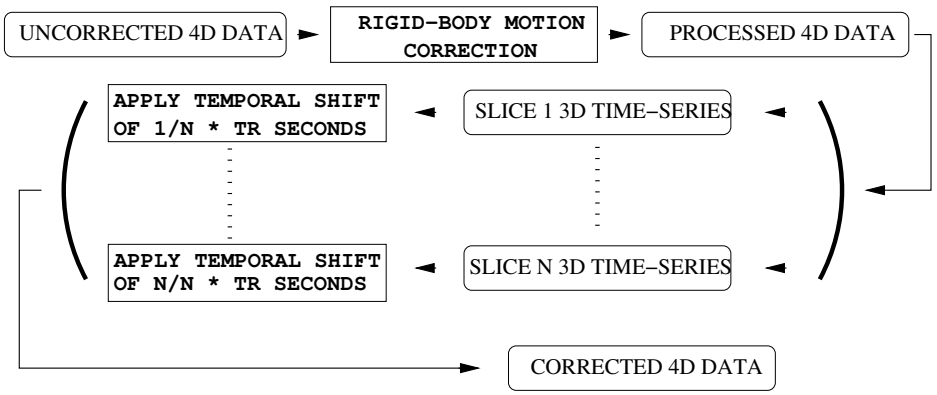


Figure 1. Schematic of the separate application of spatial and temporal realignment of FMRI data. Note that in this arrangement, motion correction is applied before slice-timing. The reverse ordering is discussed in the text.

Performing slice-timing correction means, ideally, that the image reflects a true 'snapshot' of the object at a discrete point in time, rather than as a volume of slices acquired sequentially. The slice-timing-corrected data should more

*Corresponding author: Dr P R Bannister, prb@robots.ox.ac.uk.

accurately reflect the spatio-temporal relationship between voxels throughout the volume. Because the EPI data-sets under consideration have typically been acquired slice-by-slice rather than by using a volumetric approach, the temporal distribution of the slices within each volume must be taken into account if correct and meaningful conclusions are to be drawn about the spatio-temporal relationships within the data.

While the need for slice-timing is widely acknowledged, these corrections are not always performed in practice on FMRI data. Applying the two corrections separately is convenient, not least because it facilitates the use of existing tools for rigid-body motion correction along with separate temporal interpolation of each voxel time-course to re-shift slice-timings. There are fundamental errors in the assumptions underlying this distinction, however, regardless of the order in which the two steps are performed. Clearly, if no subject motion has occurred, it is sufficient simply to apply slice-timing correction as a series of temporal interpolations over each voxel time-course in turn, where the amount of shift is proportional to the temporal offset associated with the slice containing the voxel being considered, shown in Figure 1. A complete lack of subject motion is unlikely to occur in real data, however, so the interaction between motion and acquisition delays must be modelled in order to correct fully for the resulting artefacts in the data.

Assuming that motion correction is carried out before any temporal corrections, data which may not correspond to acquisition at a consistent point in time will be co-registered. If slice-timing correction is applied after the initial realignment, the corrected images will contain data from several discrete sample times within individual slices. This is because, in the general case of through-plane motion, spatial registration will realign the data so that intensity values from individual slices in scanner space are distributed across several slice locations in the corrected data. Specifically, if rigid-body realignment is performed, subsequent slice-timing will make the incorrect assumption that data within individual slices will have been acquired at the same time-point. In this situation, it is necessary to keep a record of the slice in which the data were originally acquired and then apply the appropriate timing correction, a step which is usually omitted.

It might therefore seem obvious that temporal re-sampling should be carried out before motion correction. An obstacle to such a re-sampling is that in order to carry out slice-timing correction by temporal interpolation of a particular voxel, the time-course of that voxel must be known. If the subject has moved, there is no guarantee that a voxel in object space will be in constant alignment with a voxel in scanner coordinates. This creates a cyclic problem where motion correction is needed in order to determine slice-timing before motion correction. The purpose of the work described in the remainder of this paper is to demonstrate that the use of separate spatial and temporal algorithms to solve a spatio-temporal realignment problem cannot be considered a suitable approximation to the true solution.

3 Experimental Results

A number of artificial data-sets were generated which contained a range of intra-slice movements. This was done by applying the transformations describing the motion to a high-resolution (256x256x128 voxel) 20 volume EPI time-series containing no prior motion or activation. The data was then down-sampled to a more typical size of 64x64x21 voxels per volume.

Four possible corrections were applied: RAW (no correction), MC (motion correction using MCFLIRT [1]), MC+ST (MCFLIRT followed by temporal sinc interpolation) and ST+MC (sinc interpolation followed by MCFLIRT). After correction, the level of correction was characterised by examining the median average residual error across all the voxel time-courses within the time-series. The error is computed by comparing the intensity values in the corrected data to those in the original (unperturbed) images. In the case of a perfect correction, this value should be zero.

In the first case, two data-sets containing either nodding (rotation around a central x-axis) or shaking (rotation around a central z-axis) were evaluated. The range of motions were +1 , +2, +1, -1, -2, -1 , +1 ... per volume with the motion applied incrementally over each slice. For example, for the first slice of the first volume, a rotation of 0.0476 was added, for the second slice 0.0952 and so on. The results are presented in Figures 2 and 3. For the shaking motion, motion correction alone was able to slightly reduce the error but additional slice-timing, both prior and post MCFLIRT, made this correction worse. In the case of the nodding data, all three correction approaches led to a worse error than no correction at all.

A second group of data-sets were then generated which contained a range of either nodding or shaking motions

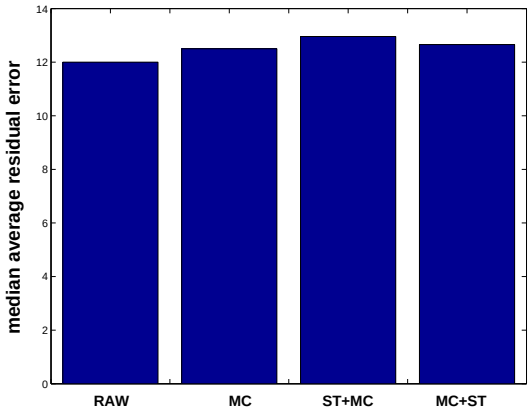


Figure 2. Comparison of relative accuracies, measured as median average residual variance, of separate spatial and temporal corrections when applied to data known to contain inter-slice movement within individual volumes. *From left to right*: Uncorrected data (RAW), MCFLIRT motion correction (MC), slice-timing correction then MCFLIRT (ST+MC), MCFLIRT then slice-timing correction (MC+ST).

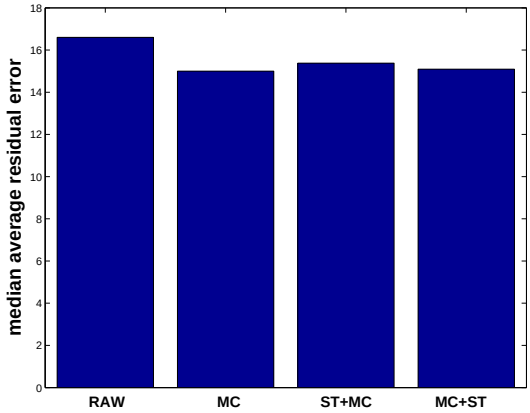


Figure 3. Comparison of relative correction errors, measured as median average residual variance, of separate spatial and temporal schemes when applied to data-sets which are known to contain only intra-slice movement within individual volumes. *From left to right:* Uncorrected data (RAW), MCFLIRT motion correction (MC), slice-timing correction then MCFLIRT (ST+MC), MCFLIRT then slice-timing correction (MC+ST)

applied as +M/-M/+M/-M etc. where M ranged from 0.1 up to 2 between data-sets. In the case of this regular, repetitive motion, slice-timing will have the effect of averaging the intensity values across time-points thus giving the impression of a good correction, Even so, it is not possible to predict whether the best correction with occur if the slice-timing step is placed before the motion correction, or after. The magnitude of the error is still around 50% of the original value (relating to an absolute error of several millimetres in the images). The results, shown in figures 4 and 5, depict this erratic performance. The plots also reveal that without the averaging effect of temporal interpolation, motion correction alone will still lead to a degradation of the images.

4 Discussion

It has been shown that the convenience of using separate slice-timing and motion correction is out-weighed by the adverse effect that this has on FMRI data exhibiting very simple and common types of intra-slice motion. In order to accurately compensate for these effects, an integrated approach to spatio-temporal re-alignment has been developed [2] which can accurately model the motion at an individual slice resolution.

Acknowledgements

The authors acknowledge the financial support of the EPSRC and MRC.

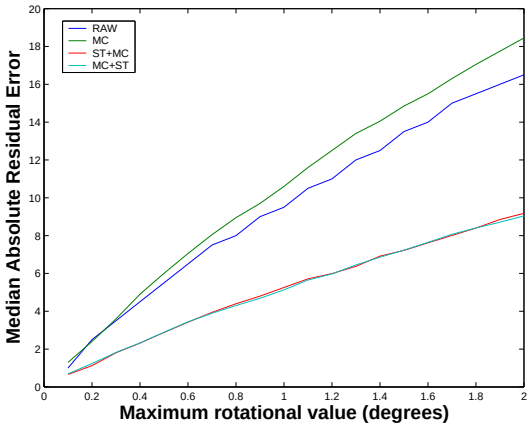


Figure 4. Plots of the median average residual variance for a shaking motion design show the effect of correction with separate slice-timing and rigid-body motion correction on images containing very low levels of motion. Notation corresponds to uncorrected data (RAW), MCFLIRT motion correction (MC), slice-timing correction then MCFLIRT (ST+MC), MCFLIRT then slice-timing correction (MC+ST).

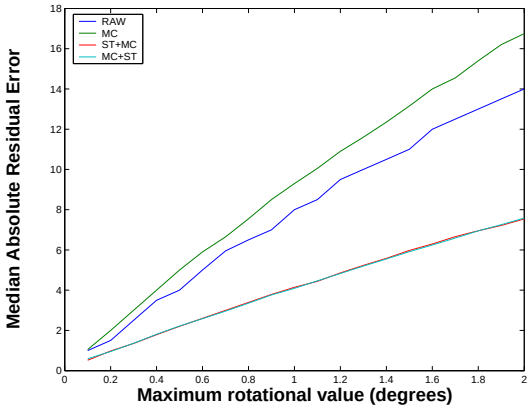


Figure 5. Plots of the median average residual variance for a nodding motion design show the effect of correction with separate slice-timing and rigid-body motion correction on images containing very low levels of motion. Notation corresponds to Uncorrected data (RAW), MCFLIRT motion correction (MC), slice-timing correction then MCFLIRT (ST+MC), MCFLIRT then slice-timing correction (MC+ST).

References

1. M. Jenkinson, P. Bannister, J. Brady et al. "Improved optimisation for the robust and accurate linear registration and motion correction of brain images." *Neuroimage* **17(2)**, pp. 825–841, 2002.
2. P. Bannister, J. Brady & M. Jenkinson. "Non-rigid motion estimation and spatio-temporal realignment in FMRI." In *3rd IEEE Workshop on Articulated and Non-rigid Motion, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2004.

Volumetric analysis of brain MRI using Geometric Deformable Models

[1]Wei Wei Lee, [2] Iain Richardson

[1&2] School of Engineering, The Robert Gordon University, Aberdeen, U.K

Email: prs.lee@rgu.ac.uk

Abstract: Repeated serial Magnetic Resonance Imaging (MRI) scans allow physicians to non-invasively examine the human brain and detect changes in brain structure over a period of time. Comparison of global and/or regional brain volume between repeated scans may be used as indicators or surrogate markers of brain pathology. Geometric Deformable Models are identified as a promising approach, since the model has the potential to accurately analyse brain volumetric changes. This paper describes the development of a Geometric Deformable Model and its application to hippocampal volume measurement.

Keyword: Volumetric analysis, Geometric Deformable Models, Magnetic Resonance Imaging

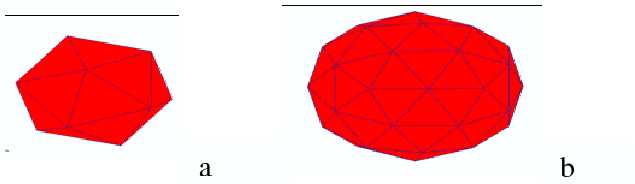
1. Introduction:

Magnetic Resonance Imaging (MRI) is widely used to image anatomical structures. Due to improvements in computer processing capability and improved imaging technologies, volumetric analysis of MRI datasets is becoming feasible. In clinical practice, simple visual inspection of MRI scans is often insufficient to determine the degree of progression or regression of pathology or to analyse the effect of treatment. Repeated serial MRI scans offers the opportunity for longitudinal studies, in which a previous scan of the patient becomes the reference point for comparison with later scans. An accurate and robust volumetric segmentation of all or part of the human brain, based on Magnetic Resonance Imaging (MRI) scans, could be use to aid early clinical diagnosis of neurodegenerative diseases. For this challenging task, Geometric deformable models (GDM) seem to be a promising approach to extract anatomical structure from volumetric data, since GDMs have the potential to (a) incorporate a range of a priori knowledge about the target structure and (b) automatically measure volume. GDMs have been successfully used to segment a variety of structures, for example: teeth [1], the cardiac left ventricle [2] and human vertebra [3]. This paper focuses on volumetric measurement of the hippocampus from MRI datasets by using GDM. Quantitative analysis (segmentation) of hippocampus is desirable since (for example) early stage of Alzheimer's disease is known to be linked to neuronal damage of the hippocampus. In clinical practice, segmentation of the hippocampus is typically performed manually by a radiologist. Manual segmentation methods bring many drawbacks such as high bias results, labour-intensive, and lack of consistency. Previous authors have applied GDM to the problem of segmenting the hippocampus [4], in which they still have not find the solution on solving mesh self intersection problem. Our GDM approach is different from [4] since we introduce of a new constraint that takes into account the relationship of non- neighbouring vertices in order to prevent GDM mesh self intersection.

This paper describes a volume analysis algorithm designed to be applied to standard quality T1-weighted longitudinal brain MRI scans in order to quantify hippocampus volume changes. The paper will first describe the GDM that we have developed, followed by a short explanation of how we measure volume from the GDM. After that, some initial results are presented. Finally, the results are discussed and future work is described.

2. Method Overview:

A Geometric Deformable Model (GDM) (sometimes described as a "balloon model") is a collection of polygons in space, forming a curve (in 2D) or a surface/solid (in 3D) (Figure 1). We are developing a GDM for MRI volumetric analysis, derived from the original formulation of [1], but incorporating a number of modifications. The operation of the GDM is based on constraint modelling and cost function minimization. Our GDM incorporates 5 constraints which are integrated together to form a local cost function (potential function) associated with each vertex in a 3D model. The geometric model is iteratively deformed to a position that minimizes its local cost function [1].



a b

Figure 1: Balloon model in mesh a) 20 triangles, b) 80 triangles (after deformation and 1 level of global subdivision.)

The local cost function at the current location is given as:

Local cost function (x,y,z) = *a0* Deformation Potential (x,y,z)+a1* Feature Event (x,y,z)+a2* Maintaining topology (x,y,z)+a3* Angular (x, y, z) +a4* Non-Neighbours (x, y, z)* (Equation 1)

Where a0, a1, a2, a3, a4 are individual weights for the following constraints:

i. *Deformation Potential*:
This constraint generates a "force" that tends to expand the model, analogous to an inflation force acting on a balloon. During each iterative step, every vertex tends to move in the normal direction of the local surface in order to decrease the deformation potential.

ii. *Feature Event:*
This constraint counteracts the deformation force, preventing the GDM from expanding beyond a target object boundary (for example, the boundary of an anatomical structure).

iii. *Maintaining topology:*
This constraint tends to preserve the model topology, minimizing the local curvature between each model point (i.e. each node on the 3D surface) and its neighbours.

iv. *Angular:*
This constraint is based on the sum of the angles that comprise the local surface and tends to keep the local mesh surface as smooth / as regular as possible.

v. *Non-Neighbouring Vertices Distance:*
In some circumstances the surface of a 3D GDM can self-intersect and this constraint is designed to counteract this effect. Non-neighbouring vertices (i.e. vertices unconnected to the current vertex) that are close to the current vertex will lead to a large constraint, preventing mesh surfaces from intersecting.

The model is a initialised as a structure in Euclidean 3-space isomorphic to a sphere, made up of uniformly tessellated icosahedrons. The GDM is refined through a series of iterative steps until it converges to match the desired anatomical structure. During this process, global and local resampling are applied in order to maintain a uniform density of vertices [1,5].

3. Volumetric analysis

The volume of the GDM is calculated after each stage of deformation. At each iteration, the vertices of a model face are deformed from their current positions to new positions. The projection between the current and deformed face positions can be considered as being composed of a set of tetrahedrons. The change in volume of the GDM at each iteration is calculated as the sum of the volumes of the tetrahedrons making up each face projection. This volume calculation does not depend on voxel count and is capable of sub-voxel accuracy.

4. Experimental Results:

Figure 2 illustrates the results of segmenting a dataset containing a known volume (a simple sphere). Figure 3(a) is a rendering of a manually-segmented Hippocampus from MRI volume data. The boundary of the manual segmentation was applied as the Feature constraint to our GDM (see section 3 above) (a more straightforward segmentation task than using the MRI data directly as a Feature constraint). The result of applying the GDM is shown in Figure 3(b). A comparison between the calculated volumes obtained (a) by counting voxels and (b) as an out put of our model is shown in Table 1.

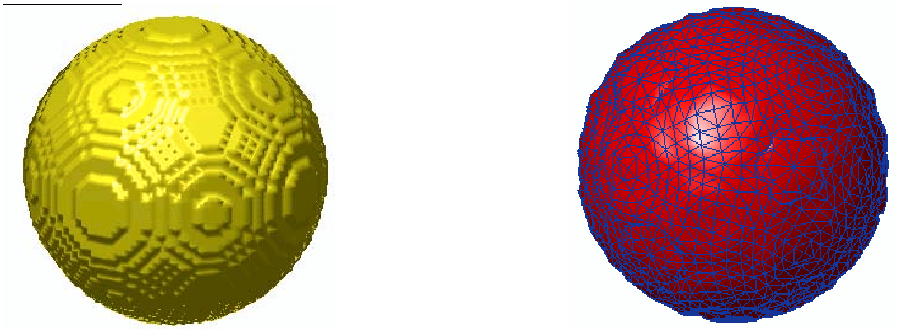


Figure 2 Segmentation of a sphere volumetric dataset: a) Rendered sphere b) GDM approximation to sphere.

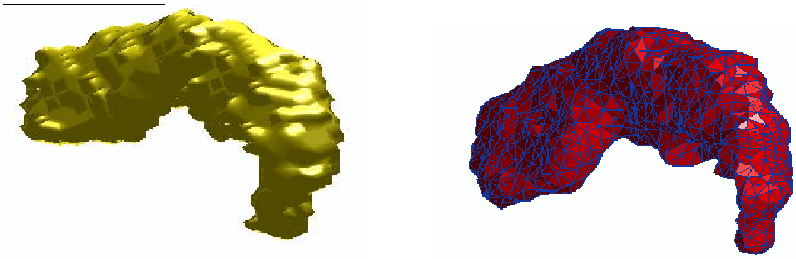


Figure 3 Segmentation of a Hippocampus: a) rendered manual segmentation b) GDM approximation to hippocampus.

Table 1

Object	Voxel count volume	Model volume
Sphere	54435	54402
Hippocampus	1659	1620.5

5. Discussion and Future work:

In section 4, we show some initial experimental results obtained from our model. Currently, our model can perform segmentation and volume analysis on regular geometric shapes and on objects that have clearly defined boundaries (for example, the manually segmented boundary shown in Figure 3). The results in Table 1 indicate that the GDM produces a volume estimate that is close to that produced by voxel counting, but is slightly smaller due to the fact that the GDM produces a smoother object surface (and hence a more accurate volumetric measurement).

The hippocampus is surrounded by different kinds of tissue and on some parts of the hippocampus boundary there is no local feature to define its edge, making automatic segmentation a difficult task. We are currently developing feature constraints to deal with the problem of automatic segmentation directly from MRI data. Our goal is to develop a model that is capable of accurately measuring the volume of structures such as hippocampus with minimal operator intervention (e.g. little or no manual segmentation).

Acknowledgements

The authors would like to thank Dr Roger Staff of the Department of Biomedical Physics, University of Aberdeen, for kindly providing sample datasets, manual segmentation data and clinical advice.

References

1. J.V. Miller. "On GDM's: Geometrically deformed models for the extraction of closed shapes from volume data." Masters thesis, Rensselaer Polytechnic Institute, Troy, New York, December 1990.
2. Cohen, L. D. a. C., I (1993). "Finite element methods for active contour models and balloons for 2-D and 3-D images." IEEE Transactions PAMI 15(11): 1131-1147.
3. Ye Duan, Liu Yang, Hong Qin, and Dimitris Samaras, " Shape Reconstruction from 3D and 2D Data Using PDE based Deformable Surface.", Proceedings of the 8th European Conference on Computer Vision (ECCV 2004), May 11-14, 2004, Prague, Czech
4. A. Ghanei, H. Soltanian-Zadeh, K. Elisevich, J.A. Fessler: "Knowledge-based deformable surface model with application to segment of brain structure in MRI," Presented at and published in the Proceedings of SPIE Medical Imaging 2001: Image Processing Conference, San Diego, CA, Feb. 17-22, 2001
5. H.Hoppe , T DeRose, T Duchamp , J. Mc Donald and W.stuetzle. "Surface reconstruction for unorganized points". SINGGAPH,1992