Josien P. W. Pluim
Boštjan Likar
Frans A. Gerritsen (Eds.)

LNCS 4057

# Biomedical Image Registration

Third International Workshop, WBIR 2006
Utrecht, The Netherlands, July 2006
Proceedings

Springer

Lecture Notes in Computer Science 4057

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

Josien P. W. Pluim
Boštjan Likar
Frans A. Gerritsen (Eds.)

# Biomedical Image Registration

Third International Workshop, WBIR 2006
Utrecht, The Netherlands, July 9-11, 2006
Proceedings

Springer

Volume Editors

Josien P. W. Pluim
University Medical Center Utrecht
Image Sciences Institute
Heidelberglaan 100, 3584CX Utrecht, The Netherlands
E-mail: josien@isi.uu.nl

Boštjan Likar
University of Ljubljana
Faculty of Electrical Engineering
Laboratory of Imaging Technologies
Tržaška 25, 1000 Ljubljana, Slovenia
E-mail: bostjan.likar@fe.uni-lj.si

Frans A. Gerritsen
Philips Medical Systems
Healthcare Informatics
P.O. Box 10.000, 5680 DA Best, The Netherlands
and
Technische Universiteit Eindhoven
Department of Biomedical Engineering
The Netherlands
E-mail: frans.gerritsen@philips.com

Library of Congress Control Number: 2006927802

CR Subject Classification (1998): I.4, I.5, H.3, J.3

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

ISSN 0302-9743
ISBN-10 3-540-35648-7 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-35648-6 Springer Berlin Heidelberg New York



Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11784012 06/3142 5 4 3 2 1 0

# Preface

The Third International Workshop on Biomedical Image Registration (WBIR) was held July 9-11, 2006, at Utrecht University, Utrecht, The Netherlands. Following the success of the first workshop (WBIR 1999), held in Bled, Slovenia, and the second workshop (WBIR 2003), held in Philadelphia, Pennsylvania, this meeting (WBIR 2006) aimed to once again gather leading researchers in the area of biomedical image registration so as to present and discuss recent developments in the field.

In modern medicine and biology, a valuable method of gathering knowledge about healthy and diseased organs, tissues, and cells is the integration of complementary information from volumetric images of these objects. Such information may be obtained by different imaging modalities, different image acquisition setups, different object preparation procedures, or by sequential image acquisition in follow-up studies or in dynamic imaging. A necessary pre-processing step for the integration of image information is image registration by which images, containing complementary information, are brought into the best possible spatial correspondence with respect to each other. Enabling combination and quantification of information about location, form and function, image registration is nowadays finding increasing use in diagnosis, treatment planning, and surgical guidance.

This year's workshop consisted of 20 oral presentations with ample time for discussions, 18 poster presentations and 2 tutorials: one addressing techniques and applications and the other numerical methods for image registration. We were delighted to welcome the participants to Utrecht and hope they found the meeting an interesting, fruitful, enjoyable and stimulating experience. For the readers unable to attend the workshop, we hope that you find these proceedings a valuable record of the scientific programme.

We would like to thank everyone who contributed to the success of this workshop: the authors for their excellent contributions, the members of the Programme Committee for their review work, promotion of the workshop and general support, the tutorial speakers for their outstanding educational contributions, the local organization staff for their precious time and diligent efforts, Philips Medical Systems for kind and generous financial support, and all the attendees for their active participation in the formal and informal discussions.

July 2006

Josien P. W. Pluim
Boštjan Likar
Frans A. Gerritsen

# Organization

## Chairs

| | |
|---|---|
| Josien P.W. Pluim | University Medical Center Utrecht, The Netherlands |
| Boštjan Likar | University of Ljubljana, Slovenia |
| Frans A. Gerritsen | Philips Medical Systems and Technische Universiteit Eindhoven, The Netherlands |

## Programme Committee

| | |
|---|---|
| Daniel Alexander | University College London, UK |
| Christian Barillot | INRIA, France |
| Michael Braun | University of Technology, Sydney, Australia |
| Gary Christensen | University of Iowa, USA |
| Albert Chung | The Hong Kong University of Science and Technology, Hong Kong |
| Louis Collins | Montreal Neurological Institute, McGill University, Canada |
| Benoit Dawant | Vanderbilt University, USA |
| Bernd Fischer | University of Luebeck, Germany |
| Mike Fitzpatrick | Vanderbilt University, USA |
| Jim Gee | University of Pennsylvania, USA |
| Alexandre Guimond | Siemens Molecular Imaging, UK |
| Mark Jenkinson | FMRIB Centre, University of Oxford, UK |
| Tianzi Jiang | National Laboratory of Pattern Recognition, China |
| Murray Loew | George Washington University, USA |
| Frederik Maes | K.U. Leuven, Belgium |
| Twan Maintz | Utrecht University, The Netherlands |
| Calvin Maurer Jr. | Accuray, Inc., USA |
| Chuck Meyer | University of Michigan, USA |
| Sébastien Ourselin | CSIRO ICT Centre, Australia |
| Xavier Pennec | INRIA, France |
| Graeme Penney | University College London, UK |
| Franjo Pernuš | University of Ljubljana, Slovenia |
| Terry Peters | Robarts Research Institute, London, Canada |
| Alexis Roche | Service Hospitalier Frederic Joliot, France |
| Torsten Rohlfing | SRI International, USA |
| Karl Rohr | University of Heidelberg and DKFZ Heidelberg, Germany |
| Daniel Rueckert | Imperial College London, UK |
| Dinggang Shen | University of Pennsylvania, USA |

| | |
|---|---|
| Oskar Škrinjar | Georgia Institute of Technology, USA |
| Colin Studholme | University of California, San Francisco, USA |
| Philippe Thévenaz | École Polytechnique Fédérale de Lausanne (EPFL), Switzerland |
| Max Viergever | University Medical Center Utrecht, The Netherlands |
| Simon Warfield | Harvard Medical School, USA |
| Jürgen Weese | Philips Research Laboratories Aachen, Germany |
| Sandy Wells | Harvard Medical School and Brigham and Women's Hospital, USA |

## Local Organization

| | |
|---|---|
| Renee Allebrandi | University Medical Center Utrecht, The Netherlands |
| Gerard van Hoorn | University Medical Center Utrecht, The Netherlands |
| Marjan Marinissen | University Medical Center Utrecht, The Netherlands |
| Marius Staring | University Medical Center Utrecht, The Netherlands |
| Jacqueline Wermers | University Medical Center Utrecht, The Netherlands |
| Darko Škerl | University of Ljubljana, Slovenia |

## Sponsoring Institution

Philips Medical Systems, The Netherlands PHILIPS

# Table of Contents

# Medical Image Registration Based on BSP and Quad-Tree Partitioning

A. Bardera, M. Feixas, I. Boada, J. Rigau, and M. Sbert

Institut d'Informàtica i Aplicacions, Universitat de Girona, Spain
{anton.bardera, miquel.feixas, imma.boada,
jaume.rigau, mateu.sbert}@udg.es

**Abstract.** This paper presents a study of image simplification techniques as a first stage to define a multiresolution registration framework. We propose here a new approach for image registration based on the partitioning of the source images in binary-space (BSP) and quad-tree structures. These partitioned images have been obtained with a maximum mutual information gain algorithm. Multimodal registration experiments with downsampled, BSP and quadtree partitioned images show an outstanding accuracy and robustness by using BSP images, since the grid effects are drastically reduced. The obtained results indicate that BSP partitioning can provide a suitable framework for multiresolution registration.

## 1 Introduction

Multimodal image registration plays an increasingly important role in medical imaging. Its objective is to find a transformation that maps two or more images, acquired using different imaging modalities, by optimizing a certain similarity measure. Among the different similarity measures that have been proposed, mutual information (MI)[2, 9] and normalized mutual information (NMI)[6] are the most commonly used since they produce satisfactory results in terms of accuracy, robustness and reliability. However, MI-based methods are very sensitive to implementation decisions, such as interpolation and optimization methods, and multiresolution strategies [4]. The latter allow us to reduce the computational cost by means of a coarse-to-fine hierarchical representation of the images. Crucial to building these hierarchies is the selection of the image simplification strategy.

The main objective of this paper is to analyze the behavior of the registration process when the source images are simplified in BSP and quad-tree structures, obtained with a maximum MI gain algorithm [5]. We will see that multimodal registration experiments based on BSP partitioned images show a remarkable accuracy and robustness, reducing substantially the grid effects compared with both regular downsampled and quad-tree images. Since experimental results demonstrate the good performance using these simplification strategies, we suggest they are an ideal strategy for defining a multiresolution framework. Such a framework can be used not only for registration purposes but also for image processing or transmission in telemedicine environments.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 1–8, 2006.

This paper is organized as follows. In Section 2, we briefly describe image registration and partitioning techniques using MI maximization. In Section 3, a new image registration framework based on partitioned images is presented. In Section 4, multimodal registration experiments show the suitability of the presented approach. Finally, our conclusions are given in Section 5.

## 2 Previous Work

In this section we review the MI definition [1] and its application to image registration [2, 9, 4, 7] and partitioning [5].

**Mutual Information.** Given two discrete random variables, $X$ and $Y$, with values in the sets $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $\mathcal{Y} = \{y_1, \ldots, y_m\}$, respectively, the MI between $X$ and $Y$ is defined as

$$I(X,Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log \frac{p_{ij}}{p_i q_j} \tag{1}$$

where $n = |\mathcal{X}|$, $m = |\mathcal{Y}|$, $p_i = Pr[X = x_i]$ and $q_j = Pr[Y = y_j]$ are the marginal probabilities and $p_{ij} = Pr[X = x_i, Y = y_j]$ is the joint probability. $I(X,Y)$ is a measure of the shared information between $X$ and $Y$. It can also be expressed as $I(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$, where $H(X)$ and $H(Y)$ are the marginal entropies, and $H(X|Y)$ and $H(Y|X)$ the conditional entropies [1].

A fundamental property of MI is the *data processing inequality* which can be expressed in the following way: if $X \rightarrow Y \rightarrow Z$ is a Markov chain, then

$$I(X,Y) \geq I(X,Z). \tag{2}$$

This result demonstrates that no processing of $Y$, deterministic or random, can increase the information that $Y$ contains about $X$ [1].

**MI-based Image Registration.** The most successful automatic image registration methods are based on MI, which is a measure of the dependence between two images. The registration of two images is represented by an information channel $X \rightarrow Y$, where the random variables $X$ and $Y$ represent the images. Their marginal probability distributions, $\{p_i\}$ and $\{q_j\}$, and the joint probability distribution, $\{p_{ij}\}$, are obtained by simple normalization of the marginal and joint intensity histograms of the overlapping areas of both images [2]. The registration method based on the maximization of MI, almost simultaneously introduced by Maes et al. [2] and Viola et al. [9], is based on the conjecture that the correct registration corresponds to the maximum $MI$ between the overlapping areas of the two images. Later, Studholme et al. [6] proposed a normalization of MI defined by

$$NMI(X,Y) = \frac{H(X) + H(Y)}{H(X,Y)} = 1 + \frac{I(X,Y)}{H(X,Y)}, \tag{3}$$

which is more robust than MI, due to its greater independence of the overlap area.

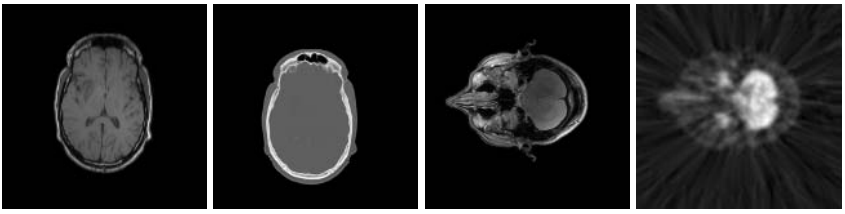
The behavior of the MI-based method depends on the implementation decisions. Thus, for instance, it is especially sensitive to the interpolator and optimizator chosen or the binning and multiresolution strategies [4]. Generally the grid points of the transformed image do not coincide with the grid points of the reference image. Thus, the selection of an interpolator is required. Although there are different interpolators, all of them introduce artifacts due to the error patterns caused by the grid regularity [7]. On the other hand, the simple computation of an MI-based similarity measure by sampling the images on a regular grid leads to undesired artifacts, called *grid effects* [8].

**MI-Based Partitioning Algorithm.** An MI-based algorithm was presented by Rigau et al. [5] to partition an image. Given an image with $N$ pixels and an intensity histogram with $n_i$ pixels in bin $i$, a discrete information channel $X \rightarrow Y$ is defined, where $X$ represents the bins of the histogram, with marginal probability distribution $\{p_i\} = \{\frac{n_i}{N}\}$, and $Y$ the image partitioned into pixels, with uniform distribution $\{q_j\} = \{\frac{1}{N}\}$. The conditional probability distribution $\{p_{j|i}\}$ of this channel is defined as the transition probability from bin $i$ of the histogram to pixel $j$ of the image, and vice versa for $\{p_{i|j}\}$. This channel fulfills that $I(X,Y) = H(X)$ since, knowing the output (pixel), there is no uncertainty about the input bin of the histogram. From the data processing inequality (2), any clustering or quantization over $X$ or $Y$, respectively represented by $\widehat{X}$ and $\widehat{Y}$, will reduce the MI of the channel. Thus, $I(X,Y) \geq I(X,\widehat{Y})$ and $I(X,Y) \geq I(\widehat{X},Y)$.

From the above reasonings, a pixel clustering algorithm which minimizes the loss of MI could be proposed. However, its high cost suggests adopting the contrary strategy, where the full image is taken as the unique initial partition and is progressively subdivided according to the maximum MI gain for each partitioning step. This algorithm is a greedy top-down procedure which partitions an image in quasi-homogeneous regions. This method can be visualized from equation $H(X) = I(X,\widehat{Y}) + H(X|\widehat{Y})$, where the acquisition of information increases $I(X,\widehat{Y})$ and decreases $H(X|\widehat{Y})$, producing a reduction of uncertainty due to the equalization of the regions. Different stopping criteria can be used. For more details, see [5].

## 3 Registration from Partitioned Images

Registration aims to find a transformation which maps two or more images by optimizing certain similarity measure. Multiresolution and multisampling strategies can be used to reduce its computational cost by means of a coarse-to-fine hierarchical strategy which starts with the reference and floating images on a coarser resolution. The estimates of the correspondence or parameters of the mapping functions while going up to finer resolutions are progressively improved. At every level they considerably decrease the search space and thus save computational time. In particular, downsampling techniques cause a great acceleration of the registration process [4].

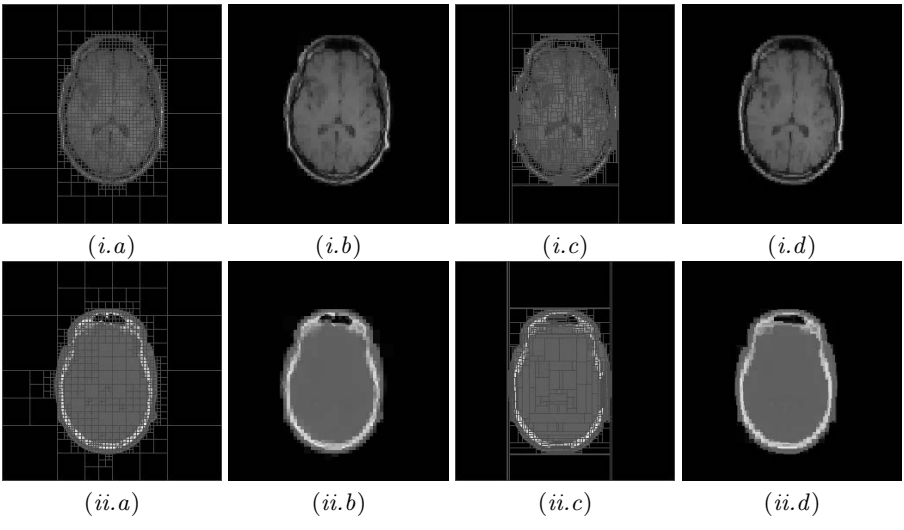


($a$) MR ($b$) CT ($c$) MR ($d$) PET

**Fig. 1.** Test images from the Vanderbilt database [3]

Obviously, a good strategy to speed-up the registration process could be to use simplified images instead of the original ones. Our proposal is to register the images obtained with the MI-based partitioning algorithm presented in Sec. 2. These images contain a high information level for a reduced number of regions. This proposal is a first approximation for considering the benefits of a multiresolution approach which would consist in the interplay of the different resolutions of both images to accelerate registration. At each registration level, the best suited resolution for each image would be selected. Crucial to developing this multiresolution framework is the selection of the simplification strategy that has to be applied to simplify images. In this paper, we investigate two subdivision techniques, BSP and quadtree, to determine which provides better results.

To carry out this study, we propose a two step registration process. In the first step, the original images are progressively partitioned with vertical or horizontal



($i.a$) ($i.b$) ($i.c$) ($i.d$)

($ii.a$) ($ii.b$) ($ii.c$) ($ii.d$)

**Fig. 2.** ($i$) MR and ($ii$) CT images obtained from Fig. 1($a$-$b$). ($a$) Quad-tree partitions with $MIR = 0.7$, ($b$) quad-tree simplified images, ($c$) BSP partitions with $MIR = 0.7$, and ($d$) BSP simplified images.

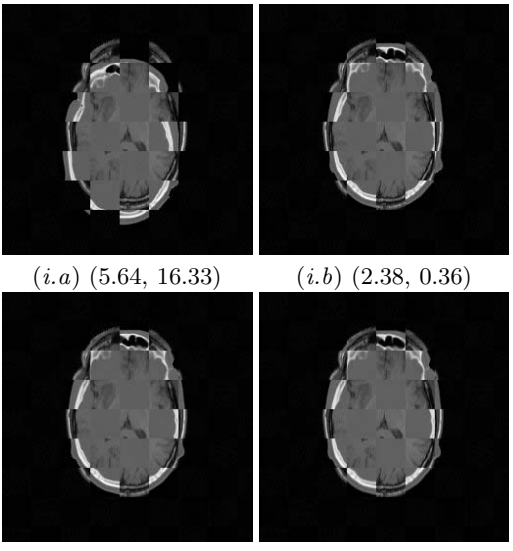
**Table 1.** Percentage of regions obtained with respect to the initial number of pixels corresponding to MR and CT original images of Fig. 1(*a-b*) and for a given $MIR$

| $MIR$ | MR | | CT | |
|---|---|---|---|---|
| | BSP | quad-tree | BSP | quad-tree |
| 0.5 | 0.25 | 0.40 | 0.06 | 0.13 |
| 0.6 | 0.81 | 1.18 | 0.21 | 0.39 |
| 0.7 | 2.21 | 3.16 | 0.77 | 1.28 |
| 0.8 | 5.28 | 6.56 | 2.73 | 3.87 |
| 0.9 | 11.88 | 16.05 | 7.98 | 11.48 |

lines (BSP) or with a quad-tree structure. In both cases, an MI ratio given by $MIR(X,\widehat{Y}) = \frac{I(X,\widehat{Y})}{H(X)}$ is used as a stopping criterion. This ratio is a measure of the simplification quality.

In Fig. 2 we illustrate the behaviour of this partitioning step applying it to the 2D MR-CT pair of images (Fig. 1(*a-b*)). In Fig. 2(a,c) we show for each original image the partitioning lines of the quad-tree and BSP structures and in Fig. 2(b,d) the corresponding *simplified* images obtained by averaging for each region the intensity of its pixels. We also collect in Table 1 the percentage of regions obtained with the simplification with respect to the initial number of pixels corresponding to the original MR and CT images. Note that a big gain of MI is obtained with a relative low number of partitions. Thus, for instance, in the CT case, a 70% of MI ($MIR = 0.7$) is obtained with approximately 1% of the maximum number of partitions (number of pixels of the source image). Observe that less partitions are needed in the CT image to extract the same $MIR$ than in the MR image. This is due to the fact that the higher the image homogeneity, the higher the degree of simplification. In this example, the CT image is more homogeneous than the MR image.

In the second step of the process, the previously partitioned images are registered using the NMI metric and the Powell's algorithm as optimizer. To illustrate the feasibility of this proposal, we have registered simplified images of the MR-CT of Fig. 1(*a-b*), considering first an $MIR$ of 0.6 and then an $MIR$ of 0.7. The registration results are shown in Fig. 3, where, respectively, $(a)$ and $(b)$ correspond to $MIR = 0.6$ and $MIR = 0.7$, and $(i)$ and $(ii)$ to the quad-tree and BSP partitioned images. In this figure, to illustrate better the obtained results, we apply the transformation obtained from the registration of the simplified images to the original ones. In addition, for each one of these images we compute the translational error ($t_x$,$t_y$). We consider the registration result of the original images without any partitioning process as being correct, so this error measures the deviation in $x$ and $y$ translation between the transformation corresponding to the correct registration and the evaluated one. In all the cases, the rotational error has been omitted due to its insignificant value. Observe that BSP images with $MIR = 0.6$ (Fig. 3($ii.a$)) achieve a lower error than quad-tree images with $MIR = 0.7$ (Fig. 3($i.b$)). This demonstrates that better results are obtained with the registration of the BSP partitioned images.



($i.a$) (5.64, 16.33) ($i.b$) (2.38, 0.36)

($ii.a$) (0.36, 0.71) ($ii.b$) (0.36, 0.17)

**Fig. 3.** Registration for the MR-CT pair of Fig. 1($a$-$b$). ($i$) Quad-tree and ($ii$) BSP subdivision methods for ($a$) $MIR = 0.6$ and ($b$) $MIR = 0.7$. The translational error ($t_x$,$t_y$) is shown for each registration.

## 4 Results and Discussion

In order to evaluate more accurately the performance of the registration of MI-based partitioned images, experiments on MR-CT (Fig. 1($a$-$b$)) and MR-PET (Fig. 1c-d) images are presented. In these experiments, the corresponding pair of images have the same degree of simplification, i.e., an MR quad-tree (or BSP) image with $MIR = 0.7$ is registered with a CT quad-tree (or BSP) with the same $MIR$. These results are compared with regular downsampled images.

In Fig. 4 the results of our experiments are presented. The behavior of the NMI measure is analyzed moving the floating image one pixel at each step through the X axis from -100 to 100 pixel units around the origin. No interpolation artifacts appear since there is no pixel interpolation. In all the plots, the bottom curve corresponds to the NMI registration of the source images. The MR-CT and MR-PET results are shown in the first ($i$) and second ($ii$) rows, respectively. In Fig. 4 ($a$), we illustrate the NMI measure obtained with different downsamplings of the original images. From bottom to top, the NMI curves correspond to downsampling of 2×2, 4×4, 8×8 and 16×16 pixels, respectively. Note that, high artifacts appear at every $n$ pixels coinciding with the downsampling factor. In Fig. 4($b$-$c$), we illustrate the NMI values for the quad-tree and BSP partitioned images, respectively. Each curve corresponds to a different degree of simplification. From bottom to top, $MIR$ ranges from 0.9 to 0.5. Observe in Fig. 4($b$) that the quad-tree partition also produces correlation artifacts due to the regularity

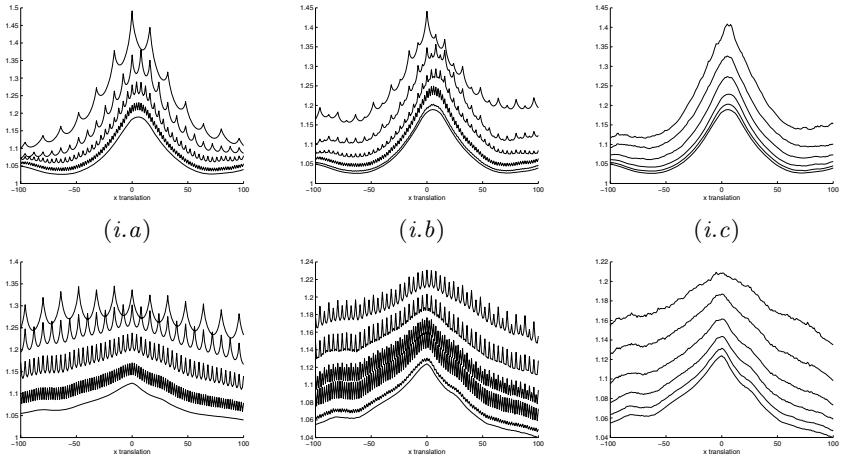



**Fig. 4.** ($i$) MR-CT (Fig. 1($a$-$b$)) and ($ii$) MR-PET (Fig. 1($c$-$d$)) registration results corresponding to ($a$) downsampled, ($b$) quad-tree, and ($c$) BSP images. The horizontal axis represents the slice translation on the X-axis (in pixels) and the vertical axis the value of NMI. For each plot, the NMI measure for different degrees of downsampling ($a$) and simplification ($b$-$c$) of the images are shown.

of its partitions. However, these artifacts are slightly reduced with respect to the downsampling case, since, although the registered images have the same degree of simplification, the number and the position of the generated quad-tree partitions are not the same.

Finally, in Fig. 4($c$) we analyze the BSP partition. In this case, the grid artifacts are nearly completely eliminated since neither the position nor the number of partitions of the images coincide. Registration is more robust since the probability of finding a local maximum is lower as it is shown by the smoothness of BSP plots. Taking into account that the perfect registration is given by the maximum bottom curve, observe the high accuracy, i.e., the coincidence of the curve maxima, of the registration reached with the BSP images. For instance, an accurate registration is achieved with $MIR = 0.7$, which represents an approximate reduction of 99% of the original number of pixels.

Experiments with the MR-PET images shown in Fig. 4($ii$) behave similarly to the MR-CT case in Fig. 4($i$). In both cases, the BSP simplification scheme behaves considerably better than both quad-tree simplification and downsampled images in terms of the reduction of grid artifacts. From these experiments we can conclude that the BSP approach is more robust and accurate.

## 5 Conclusions and Future Work

In this paper, we have presented a new technique for image registration based on the partitioning of the source images. Thepartitioning algorithm relies on the

maximization of the mutual information gain for each refinement decision. The presented method is a first step towards a full multiresolution registration approach. Two alternatives (binary space partition and quad-tree simplifications) have been analyzed and compared with a usual regular downsampling technique. The quality of the subdivision has been investigated in terms of the efficiency in registration. Results have shown the superior quality of the BSP subdivision, which allows a smoother registering. The BSP approach performs also better than regular downsampling. The next step in our research will consist in developing a multiresolution framework using the BSP subdivision.

## Acknowledgments

The images used in our experiments were provided as part of the project, "Evaluation of Retrospective Image Registration", National Institutes of Health, Project Number 1 R01 NS33926-01, Principal Investigator Prof. J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN. This project has been funded in part with grant numbers TIN2004-08065-C02-02, TIN2004-07451-C03-01 and 2001-SGR-00296.

## References

1. Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* Wiley Series in Telecommunications, 1991.
2. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.
3. National Institutes of Health. *Retrospective Image Registration Evaluation.* Vanderbilt University, Nashville (TN), USA, 2003. Project Number 8R01EB002124-03, Principal Investigator J. Michael Fitzpatrick.
4. Josien P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22:986–1004, 2003.
5. J. Rigau, M. Feixas, M. Sbert, A. Bardera, and I. Boada. Medical image segmentation based on mutual information maximization. In *International Conference on Medical Image Computing and Computed Assisted Intervention (MICCAI 2004), Proceedings*, Rennes-Saint Malo, France, September 2004.
6. Colin Studholme. *Measures of 3D Medical Image Alignment.* PhD thesis, University of London, London, UK, August 1997.
7. J. Tsao. Interpolation artifacts in multimodal image registration based on maximization of mutual information. *IEEE Transactions on Medical Imaging*, 22:854–864, 2003.
8. M. Unser and P. Thévenaz. Stochastic sampling for computing the mutual information of two images. In *Proceedings of the 5th International Workshop on Sampling Theory and Applications (SampTA'03)*, pages 102–109, Strobl, Austria, May 2003.
9. Paul A. Viola. *Alignment by Maximization of Mutual Information.* PhD thesis, Massachusetts Institute of Technology, Massachusetts (MA), USA, 1995.

# A Bayesian Cost Function Applied to Model-Based Registration of Sub-cortical Brain Structures

Brian Patenaude, Stephen Smith, and Mark Jenkinson

FMRIB Centre, University of Oxford

**Abstract.** Morphometric analysis and anatomical correspondence across MR images is important in understanding neurological diseases as well as brain function. By registering shape models to unseen data, we will be able to segment the brain into its sub-cortical regions. A Bayesian cost function was derived for this purpose and serves to minimize the residuals to a planar intensity model. The aim of this paper is to explore the properties and justify the use of the cost function. In addition to a pure residual term (similar to correlation ratio) there are three additional terms, one of which is a growth term. We show the benefit of incorporating an additional growth term into a purely residual cost function. The growth term minimizes the size of the structure in areas of high residual variance. We further show the cost function's dependence on the local intensity contrast estimate for a given structure.

## 1 Introduction

Morphometric changes in sub-cortical brain regions are associated with psychiatric disorders, neurodegenerative diseases, and aging. Furthermore, anatomical correspondence across MR images is needed to perform group analysis of functional data. Manual delineation of subcortical structures is a very time consuming task and requires considerable training. One approach to solving this problem is by registering a probabilistic brain atlas to new data [1]; a more recent approach also incorporates anisotropic Markov Random Fields and intensity priors [2]. We are proposing to solve the registration/segmentation problem by registering statistical shape models to MRI data.

A Bayesian similarity function which aims to minimize the residuals to a planar intensity model was derived to drive the registration. The aim of this paper is to investigate the cost function's properties and justify its use. Like correlation ratio, this cost function minimizes residuals, however it has three additional terms. We show that the added benefit of the full Bayesian cost function over a pure residual function is due to the addition of a growth term. The balance between the growth and residual term is governed by the local intensity contrast for a given structure. The cost function's relationship to the local intensity contrast is examined as well.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 9–17, 2006.

## 2 Methods

### 2.1 Model Building

The training set consisted of 93 manually labelled T1-weighted MRI brain scans (courtesy of the CMA in Boston). The labelling is of a high quality and their reliability and reproducibility have been documented [3, 4]. As a first step, the training data are first affine-registered to MNI152 space. The training points are then automatically assigned to the manually labelled data using deformable surfaces. Within-surface motion constraints are imposed on the deformation process to preserve point correspondence. We assume a multivariate Gaussian model, and estimate its parameters using PCA [5, 6].

Figure 1a shows the average mesh for the left putamen, pallidum, and thalamus (three sub-cortical brain structures), which are used to initialize the registration. The transformations are applied directly to the model surface meshes, which are then converted into image space for evaluation. The conversion to image space is discussed in more detail in the following section. The deformations are limited to linear combinations of the modes of variation, and are proportional to the cost-gradient in the direction of the modes of variation. Figure 1b shows the first three modes of variation for the left putamen.

### 2.2 Cost Function

Relating the surfaces to an MRI image is done using a Bayesian similarity function that was derived specifically for this purpose. It is expressed in terms of the posterior probability of a transformation $T$ (the deformation of the surface models) given the observed MR intensity data, $Y$, and the statistical shape model, $S$. In its negative log-likelihood form it acts as a cost function and has the form:

$$F_B = -\log\left(p(T|Y,S)\right) \propto -\log\left(p(T)\right) + \frac{1}{2}\log\left|\det\left(G_{in}^T G_{in}\right)\right| \\ -\log\left(\Gamma\left(\frac{N_{eff}}{2}-1\right)\right) + \frac{N_{eff}}{2}\log\left(\pi C^{-2} Y^T R Y\right) \quad (1)$$

where $p(T)$ is the prior probability of a transformation (based on the statistical shape model), $G_{in}$ is the image generator matrix (whose columns are reshaped model intensity images – see below), $N_{eff}$ is the effective number of voxels in the shapes of interest (degrees of freedom), $C$ is an estimate of the local intensity contrast, $Y^T RY$ is the residual variance in the area of interest, where $R$ is a residual forming matrix. This similarity function (described more fully in [7, 8]) is based on the principle of an image generation function that relates the surfaces to images with voxel intensities. By fitting a model to the image intensities within the mesh region, an intensity image is generated according to the parameter estimates of the model. The particular form of image generation function chosen is one that allows the voxels within an image to have a constant intensity plus three spatially-linear gradient terms. That is, a planar fit in intensity is done within the voxels bounded by each surface (with appropriate allowance for partial

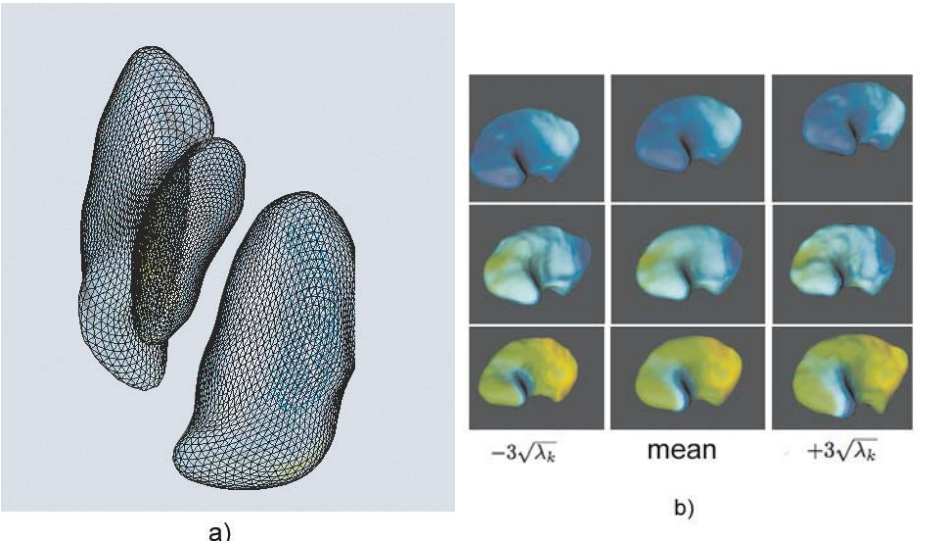



**Fig. 1.** a) Average mesh for left putamen, pallidum, and thalamus. b) First three modes of variation of left putamen. The colours are proportional to the normalized magnitude of the mode vector at each vertex. Dark blue corresponds to the largest magnitude, and transitions to green, yellow, then red, with decreasing magnitude.

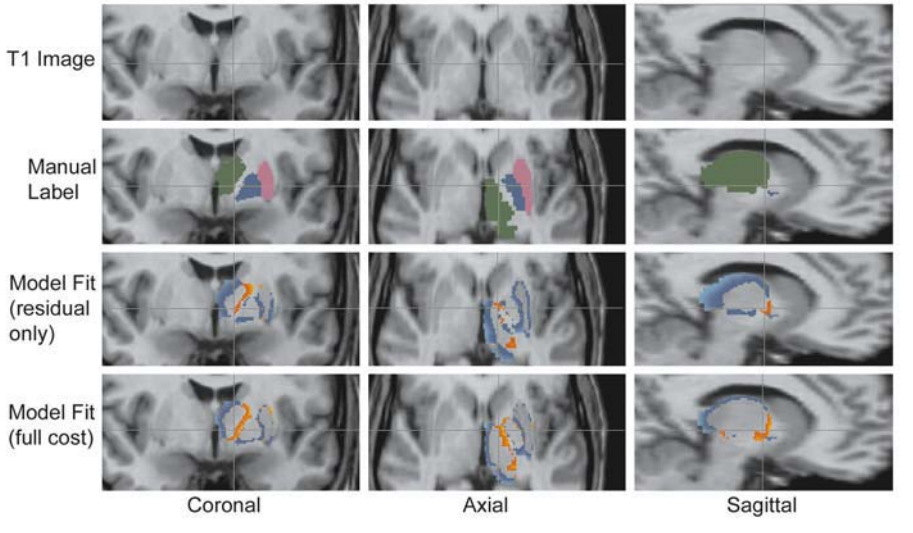



**Fig. 2.** T1 weighted image, FFND (blue scale) and FFPD (red scale) maps for subject 24 using the full and residual only cost function. There was a reduction in FD of 0.32, 0.24 and 1.36 for the left putamen, pallidum and thalamus respectively.

volume effects). The best planar fit is then subtracted from the real data leaving the residuals ($Y^T RY$). This is the same principle that the Correlation Ratio uses, except that for the Correlation Ratio only a mean value is allowed in the fitting. Here we add the linear terms in order to account for bias field effects and slow changes in tissue densities across structures.

In forming the posterior probability, prior probabilities are required for the transformation, $p(T)$ and for the intensities in regions outside the shapes of interest. The form we take for this transformation prior, $p(T)$, is a multivariate, truncated uniform distribution, where we use the statistical shape model to set the truncation limits at $\pm 3\sqrt{\lambda_k}$ (and $\lambda_k$ is the variance of the $kth$ mode in the PCA on the training shapes). For the prior on the intensities at the borders of the shape we use a uniform density with truncation proportional to an empirically derived estimate for the intensity contrast of that shape, $C$.

### 2.3 Testing

**Methods.** Models have been created for 17 sub-cortical structures, however a subset of three structures were used for testing. The left putamen, pallidum and thalamus were chosen as they span several shapes, sizes, and boundary contrasts. The putamen has good contrast with the surrounding structures, despite weaker contrast where it borders the pallidum. The putamen may also be proximal to cortical sulci which can be problematic due similar intensity ranges. Medially, the pallidum intensity values transition smoothly into the white matter, which results in weak contrast. Medially, the thalamus has high contrast with the lateral ventricles, however suffers lateral blending into the white matter. The thalamus also poses the challenge of having intensity inhomogeneities across its sub-regions; this is due to changes in the grey matter density.

Our investigation of the full (Bayesian) cost function, $F_B$ is done by comparing the registration under the full versus simple form of the cost function. The simple form is a purely residual-based cost function similar to the Correlation Ratio, although slightly different to the fourth term in the full cost function. The purely residual-based cost function, which is simpler and more intuitive, is expressed as:

$$F_R \propto -\log\left(\frac{\pi C^{-2} Y^T RY}{N_{eff}}\right) \tag{2}$$

For practical implementation/optimization purposes the probability density function of the shape model is truncated at $\pm 3\sqrt{\lambda_k}$ ($\lambda_k$ is the variance along the $k^{th}$ eigen-vector) and the value of $C$ was set to 28%, 22%, and 32% of the full intensity range for the left putamen, pallidum, and thalamus respectively. The $C$ parameter adjusts the balance between the residual and the growth term. We refer to the third term in the full cost function as the growth term; this behaves such that as a structure increases in size, the cost decreases. The degree to which this impacts the registration is governed by the $C$ parameter. The registration truncated the number of modes at 25, which explained 97.3%, 97.0%, 96.4% of the variation in the left putamen, pallidum and thalamus respectively. The multiple modes of variation were searched simultaneously using a conjugate-gradient

method. The registrations were performed on T1-weighted MRI images, each image had corresponding manual labels (provided by the CMA), which were only used for validation. All the test data was excluded from the training data. The manual labels were treated as our gold standard.

**Evaluation.** The three measures used to compare the registration output with the manual labels are defined in equation 3, 4 and 5.

$$FFND = \frac{\sum_{a \epsilon A \cap M'} \min_{b \epsilon A \cap M} D(a,b)}{N_{manual}} \tag{3}$$

$$FFPD = \frac{\sum_{a \epsilon A' \cap M} \min_{b \epsilon A' \cap M} D(a,b)}{N_{manual}} \tag{4}$$

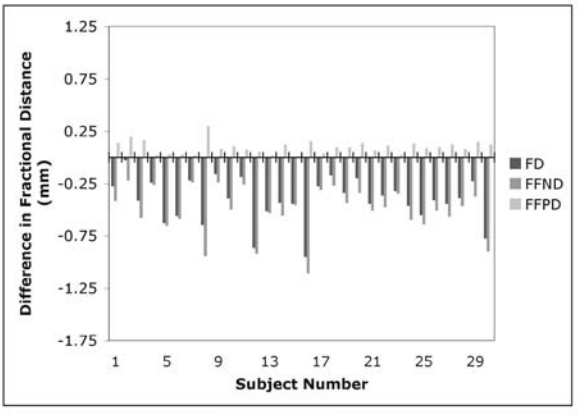
$$FD = FFND + FFPD \tag{5}$$

where FFND is the fractional false negative distance, FFPD is the fractional false positive distance and FD is the fractional distance. $A$ and $M$ are the sets of voxels assigned to a particular label by the registration and manual method respectively. $A'$ and $M'$ are the complements of the sets $A$ and $M$ respectively. $D$ is a Euclidean distance operator, and $N_{manual}$ is the volume of the manual labelled region. FFND is a distance weighted sum of the false negative voxels, normalized by the number of voxels in the manual segmentation (our gold standard). For FFND, $D$ is defined as the minimum Euclidean distance between the false negative voxel and intersection of the new label and the gold standard. FFPD is the normalized distance weighted sum of the false positive voxels. FD is the normalized distance weighted sum of all mislabelled voxels (summary statistic of total error).

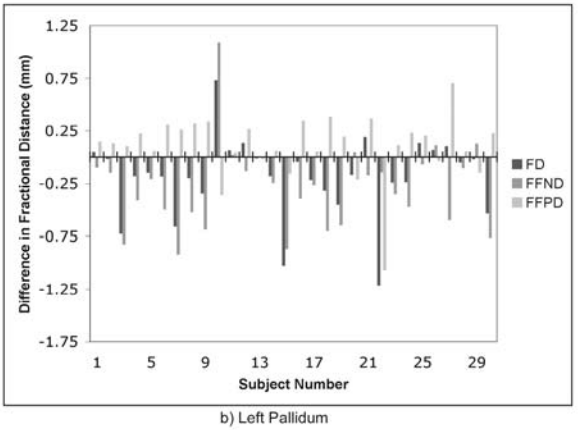
These distance-weighted sums increase sensitivity to large variations from the gold standard, thereby putting less emphasis on small dilations or erosions.

## 3 Results and Discussion

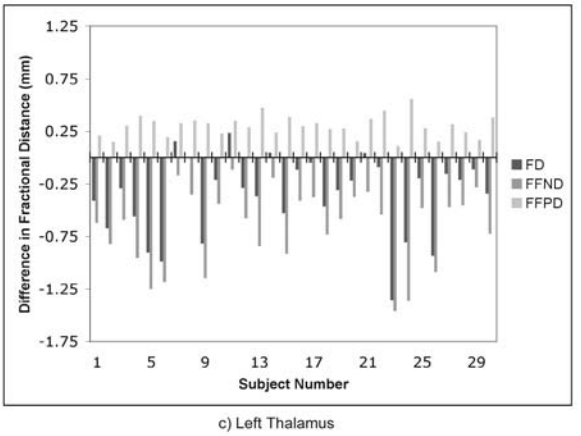
Figure 3 a, b, and c show the difference in FFND, FFPD, and FD between the full and pure residual cost function for the left putamen, pallidum and thalamus respectively. The registration was performed on 30 previously unseen images. There is a significant decrease in FD when using the full form of the cost function, suggesting a more accurate registration. The accompanying decrease in the FFND suggests that the residual form was systematically underestimating the structure, and that the addition of the growth term corrected for this. This can be seen in figure 2, where the thick blue band at the boundary of the true structure for the pure residual case signifies the under-estimation (FFND is represented by the blue scale). The additional growth term seems to allow the algorithm to




a) Left Putamen

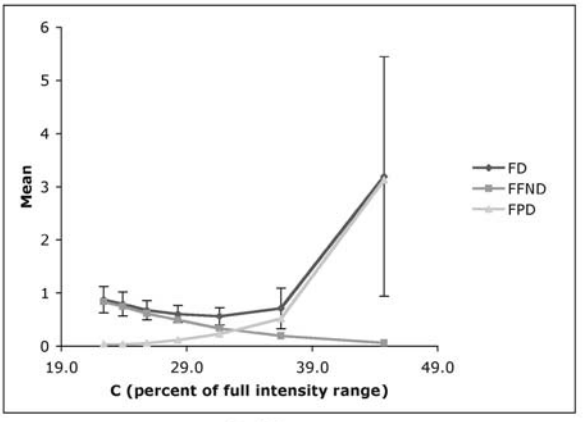



b) Left Pallidum

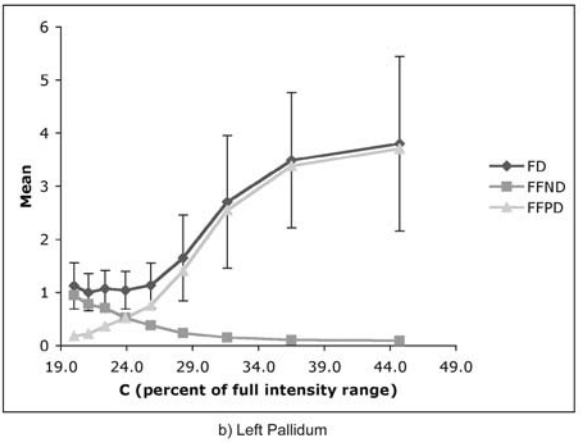



c) Left Thalamus

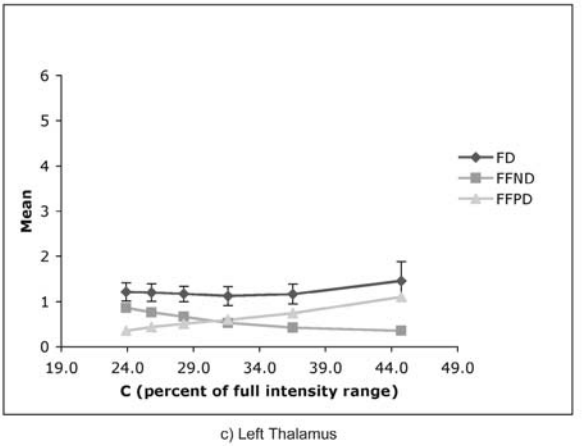
**Fig. 3.** Difference in FFND, FFPD, FD between the full and residual only cost functions for the left putamen, pallidum and thalamus. The difference is calculated such that negative values correspond to lower fractional distance in the full cost function case. FD is the summary statistic, which indicates the overall quality of the registration.




a) Left Putamen




b) Left Pallidum




c) Left Thalamus

**Fig. 4.** Mean FFND, FFPD, FD for the left putamen, pallidum and thalamus as a function of the local contrast estimate, $C$ (30 subjects). The error bars show $\pm 1$ standard deviations.

the push towards the true boundary. Though not as significantly, the improvement in FFND comes at a cost to FFPD. This is not unexpected; in the case of reduced FFND, the fit boundary is approaching the true anatomical one and is hence approaching areas of partial voluming. The cost to FFPD is evident in figure 2 in the case of the thalamus, the increase in FFPD (red band) on the lateral side of the of the thalamus is a symptom of poor contrast. In the pure residual case, the resultant structure is smaller and mostly encompassed by the true structure, hence you would expect very few false positives. The residual cost provides no reason for the model to expand to the true limits of the region, particularly if there are nonlinear inhomogeneities in the structure. The thalamus contains subregions of varying grey-matter densities (hence varying mean intensities). Our model assumes linear gradients across a structure and hence cannot properly model the abrupt changes in mean intensity.

Figure 4 shows the change in distance measures for varying local intensity contrast estimates. The FFPD is linearly increasing, however the FFND is exponentially decreasing over the region of contrasts. For each structure there is a clear minimum mean FD, and hence optimal value for $C$. The variance in FD for an appropriate value of $C$ is relatively small. For a set of images from the same scanner, $C$ may be calibrated on a single image. $C$, however, should be varied over an adequate range and sampled finely enough to cope with the large variance for poor values of $C$.

## 4 Conclusions

We have shown that the full Bayesian cost function outperforms the simple residual-only cost function. The full cost function permits structural growth in regions of low variance. This is important as it is easier to achieve a lower residual for a smaller area within a structure, and so a growth term is necessary. Indeed, this is what was observed; without the growth term the structures tended to shrink within the region of interest. The determinant term was not discussed here, as it was heavily outweighed by the residual and growth term and did not have much impact on the results. The local estimate of intensity contrast is critical to the registration performance. The performance across a range of $C$ values is well behaved, and the minimum mean FD is clear. Currently this framework for registration is modality independent, as it depends solely on residual variance after a general intensity fit (mean plus linearly varying spatial terms) per structure, and does not require any prior knowledge about absolute intensity values.

## Acknowledgments

The EPSRC IBIM grant and BBSRC for funding. IBIM collaborators for their valuable input. David Kennedy for providing us with the training data that makes this project possible.

## References


1. Collins, D., Neelin, P., Peters, T., Evans, A.: Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. Journal of Computer Assisted Tomography **18** (1994) 192–205
2. Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M.: Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron **33** (2002) 341–355
3. Kennedy, D.N., Lange, N., Makris, N., Bates, J., Meyer, J., V. S. Caviness, J.: Gyri of the human neocortex: An mri-based analysis of volume and variance. Cerebral Cortex **8** (1998) 372–384
4. Seidman, L.J., et al.: Left hippocampal volume as a vulnerable indicator for schizophrenia. Arch. Gen. Psychiatry **59** (2002)
5. Cootes, T.F., C. J. Taylor, D.H.C., Graham, J.: Active shape models—their training and application. Computer Vision and Image Understanding **61** (1995) 38–59
6. Cootes, T.F., Taylor, C.J.: Statistical models of appearance for medical image analysis and computer vision. Proc. SPIE Medical Imaging (2001)
7. Jenkinson, M., Behrens, T., Woolrich, M., Crum, W., Hill, D., Smith, S.: A bayesian similarity function for segmentation using anatomical, shape-based models. Medical Image Understanding and Analysis (MIUA) (2005)
8. Jenkinson, M.: A bayesian similarity function for segmentation using anatomical, shape-based models. Technical report, FMRIB, University of Oxford (2005)

# Automatic Inter-subject Registration of Whole Body Images

Xia Li[1], Todd E. Peterson[2], John C. Gore[2], and Benoit M. Dawant[1]

[1] Vanderbilt University, VU Station B 351662, Nashville, TN 37240-1662
{xia.li.1, benoit.dawant}@Vanderbilt.edu
http://www.vuse.vanderbilt.edu/~mip-web/
[2] VUIIS, 1161 21st Avenue South, CCC 1121 MCN, Nashville, TN 37232-2675
{todd.e.peterson, john.gore}@Vanderbilt.edu
http://vuiis.vanderbilt.edu

**Abstract.** 3D inter-subject registration of image volumes is important for tasks such as atlas-based segmentation, deriving population averages, or voxel and tensor-based morphometry. A number of methods have been proposed to tackle this problem but few of them have focused on the problem of registering whole body image volumes acquired either from humans or small animals. These image volumes typically contain a large number of articulated structures, which makes registration more difficult than the registration of head images, to which the vast majority of registration algorithms have been applied. This paper presents a new method for the automatic registration of whole body CT volumes, which consists of two steps. Skeletons and external surfaces are first brought into approximate correspondence with a robust point-based method. Transformations so obtained are refined with an intensity-based algorithm that includes spatial adaptation of the transformation's stiffness. The approach has been applied to whole body CT images of mice and to CT images of the human upper torso. We demonstrate that the approach we propose can successfully register image volumes even when these volumes are very different in size and shape or if they have been acquired with the subjects in different positions.

## 1 Introduction

Image registration is an essential tool in order to be able to follow the progression of diseases, to assess response to therapy, to compare populations, or to develop atlas-based segmentation methods. The latter involves segmenting structures in one reference volume, commonly called the atlas, and using this reference volume to segment these structures in other volumes. This necessitates being able to register the atlas to the volumes that need to be analyzed. Because it involves a number of subjects, non-rigid registration methods are required to address this problem. A number of methods and techniques have been developed over the years to achieve this; chief among them are intensity-based techniques and more specifically, methods that rely on Mutual Information (MI) [1][2]. However, most automatic methods that have been proposed have been applied to head images only. This is because head images are relatively simple compared to whole body images. Head images contain one single major identifiable structure (the cranium) as opposed to whole body images that contain many articulated

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 18 – 25, 2006.

structures (the bones). In head images the cranium surrounds the brain, therefore constraining the deformation. In whole body images, the situation is the opposite: soft tissue surrounds the bones, leading to very large inter-subject size and shape differences. All these differences make the registration of whole body images much more difficult than the registration of head images. Despite these difficulties non-rigid registration techniques for extra-cranial applications have been proposed for specific applications such as the registration of breast, abdomen, lung, or prostate images. For instance, Camara et al. [3] use a Free-Form Deformation (FFD) approach guided by a gradient vector flow combined with a grey-level MI non-linear registration algorithm for thoracic and abdominal applications. Rueckert et al. [4] also use FFD to register breast images acquired before and after contrast injection; these are image volumes acquired from the same subject. Cai et al. [5] present a validation study of CT and PET lung image registration and fusion based on the chamfer-matching method; this study also involves images acquired from the same subject.

In general, however, fully automatic inter-subject or even intra-subject registration of whole body images remains a challenge. One of the main reasons is that, in practice, non-rigid registration algorithms need to be initialized with a rigid or affine transformation. If the image volumes do not contain articulated structures, as is the case for head images, a single transformation is sufficient. If, on the other hand, these image volumes contain a number of bony structures, which are rigid but whose relative position changes from acquisition to acquisition, a single transformation is insufficient. A number of transformations need to be computed, one for each element in the articulated structure. These transformations then need to be somehow combined. This is the approach followed by Little et al. [6]. These authors present a technique designed for the intra-subject registration of head and neck images. Vertebrae are registered to each other using rigid body transformations (one for each pair of vertebrae). Transformations obtained for the vertebrae are then interpolated to produce a transformation for the entire volume. One problem with the approach is that it requires segmenting and identifying corresponding vertebrae in the image volumes. Because corresponding vertebrae are registered with rigid-body transformations, the approach is also applicable only to intra-subject registration problems. Martin-Fernandez et al. [7] propose a method, which they call articulated registration. This approach requires the labeling of landmarks to define wire models that represent the bones. A series of affine transformations are computed to register the rods, which are the elements of the wires. The final transformation for any pixel in the image is obtained as a linear combination of these elementary transformations with a weighting scheme that is inversely proportional to the distance to a specific rod. This technique has been applied to the registration of hand radiographs. Arsigny et al. [8] also propose an approach in which local rigid or affine transformations are combined. They note that simple averaging of these transformations leads to non-invertible transformations, and they propose a scheme that permits the combination of these local transformations, while producing an overall transformation that is invertible. Their method is applied to the registration of histological images. The authors comment on the fact that their method could also be used for articulated structures but do not present examples. Recently, Papademetris et al. put forth an articulated rigid registration method that is applied to the serial registration of lower-limb mouse images [9]. In this approach, each individual joint is labeled and the plane in which

the axis of rotation for each joint lies is identified. A transformation that blends piecewise rotations is then computed. The authors comment of the fact that piecewise rigid models often lead to transformations that are discontinuous at the motion boundaries, which produces folding and stretching. The approach they propose produces a transformation that is continuous at these interfaces. The authors have applied their method to the registration of lower limbs in serial mouse images. They suggest that their technique could be used to initialize an intensity-based algorithm but do not present results.

In summary, a survey of the literature shows that only a few methods have been proposed to register images including articulated structures. The general approach is to compute piecewise rigid or affine transformations and to somehow blend and combine these transformations. Unfortunately, this approach is often not practical because it requires identifying various structures in the images such as joints or individual bones. In this paper we propose a method that does not require structure labeling. This method can thus be automated, and we demonstrate its performance on small animal and human images.

## 2 Methods

There are two steps in the automatic registration method we propose. In the first step, we register only bony structures and the outside body surfaces. The transformation we compute in this first step is then used to initialize an intensity-based registration algorithm. Because our aim is to develop a fully automatic technique, we have ruled out methods that require identifying and labeling homologous structures. These methods would indeed require developing general and robust feature extraction algorithms, which is not easy to achieve. Hence, in our first step, we have chosen to rely on the robust point-based registration algorithm proposed by Chui et al. [10]. This algorithm takes as input two clouds of points and iteratively computes a correspondence between these points and the transformation that registers them, without requiring manual labeling. In addition, the two sets of points also do not need to have the same cardinality and the algorithm can deal with the problem of outliers. Correspondence is computed with the softassign algorithm proposed by Gold et al. [11]. Once correspondence is determined, a thin plate spline-based non-rigid transformation is computed to register the points. Because we use this algorithm as an initial step, the transformation it produces does not need to be extremely accurate. Point clouds in the two volumes can thus be selected in a somewhat arbitrary fashion.

In the approach we have tested so far, bone surfaces are first extracted, which can be done easily in CT images with a simple threshold. We do this in both image sets and sample the two surfaces to create the two clouds of points. Currently, we do not use any geometric feature, such as the surface curvature, to select the points. Results will show that this approach leads to acceptable results even when the skeletons are in very different positions. We then extract the external surface of the body. This is also easily achieved with an intensity threshold. As is the case for the bone surfaces, the whole body surfaces are sampled to create a second cloud of points that is added to the first one. This leads to two clouds of points, one per image volume, that typically contain 1000 to 3500 points, which are registered using the robust point-based approach of Chui et al.

The second step in our approach relies on an intensity-based registration algorithm we have proposed recently [12], which we call ABA for adaptive bases algorithm, to refine the results obtained in the first step. In this algorithm, the deformation field that registers the two images is modeled as a linear combination of radial basis functions with finite support. Coefficients for these basis functions are computed that maximize the normalized mutual information (NMI) between the images. As is often the case for non-rigid registration algorithms based on basis functions, our algorithm includes mechanisms designed to produce transformations that are topologically correct (i.e., transformations that do not lead to tearing or folding). This is done by imposing constraints on the relative value of the coefficients of adjacent basis functions. Furthermore, we compute both the forward and the backward transformations simultaneously, and we constrain these transformations to be inverses of each other. In our experience, this leads to transformations that are smooth and regular.

In our application, there are two broad categories of structures: bones and soft tissues. Because we are dealing with inter-subject registration issues, both bones and soft tissues need to be deformed (in the intra-subject registration case, individual bones can be registered with rigid-body registration methods). However, the amount of deformation typically observed for bony and soft tissue structures is very different, i.e., two livers can have vastly different shapes and sizes when the overall shape and size of individual bones vary little across subjects. This suggests using transformations whose physical properties vary spatially. These transformations should be relatively stiffer for bony structures than they are for soft tissue structures. Our algorithm allows us to do pre-cisely this. As mentioned above, regularization of the deformation field in our algorithm is obtained by imposing constraints on the relative value the coefficients associated with adjacent basis functions. In practice, we impose a threshold on the difference between the values of these coefficients. The smaller the threshold, the stiffer the transformation is. We can thus define what we call stiffness maps, which are maps that specify the value of this threshold in various regions of the image. In previous work [13], we have shown that this feature improves atlas-based segmentation results when the patient image volume contains very large ventricles or space-occupying lesions. Here, we create a simple binary stiffness map: the transformation is constrained to be stiffer over bony structures than over soft tissue structures. Results obtained when using two stiffness values, one for the bones and the other for soft tissue, improve when compared to those obtained with a single value.

## 3 Results

Our approach has been evaluated on two types of images: whole body mouse scans and upper body human scans. We used an Imtek MicroCAT II small animal scanner to generate two 512x512x512 mouse CT volumes, with a voxel resolution of $0.125 \times 0.125 \times 0.125 mm^3$. Human data sets are 512x512x184 CT volumes with a voxel resolution of $0.9375 \times 0.9375 \times 3 mm^3$. Figure 1 shows results obtained with the skeletons

of mouse volumes. The left panel shows the two skeletons in their original position. The right panel shows the same but after point-based registration. Figure 2 illustrates results obtained when both steps are applied. The left panel shows one CT slice in one volume (the source) and the right panel is the corresponding slice in the other volume (the target); note the large differences in size, shape and posture between these volumes. The middle panel shows the results we obtain when registering the source volume to the target volume. To facilitate the comparison, yellow contours of the lung have been drawn on the target image and copied on all the other ones.

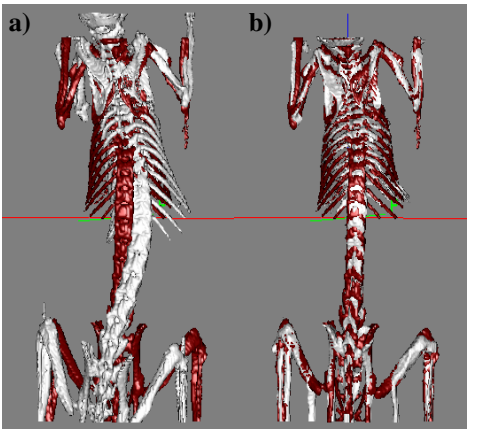



**Fig. 1.** Bony structures in two CT volumes a) before the registration and b) after the registration

Figures 3 and 4 show results we have obtained with upper torso CT images, and they illustrate the advantage of using two stiffness values. In both figures, the left panel is the source image, the right panel the target image. The second, third and fourth panels show the source volume registered to the target volume using (1) a stiff transformation, (2) a very elastic transformation, and (3) a transformation with two stiffness values. In figure 3, only bones are shown. In figure 4, the entire images are shown. When a stiff transformation is used, bones are deformed in physically-plausible ways, but soft tissues are not registered very accurately (arrows on the second panel of figure 4). When a more elastic transformation is used, bones are deformed incorrectly (regions highlighted in the third panels from the left). Using two stiffness values permits transformations to be computed that lead to satisfactory results both for the bony and soft tissue regions.

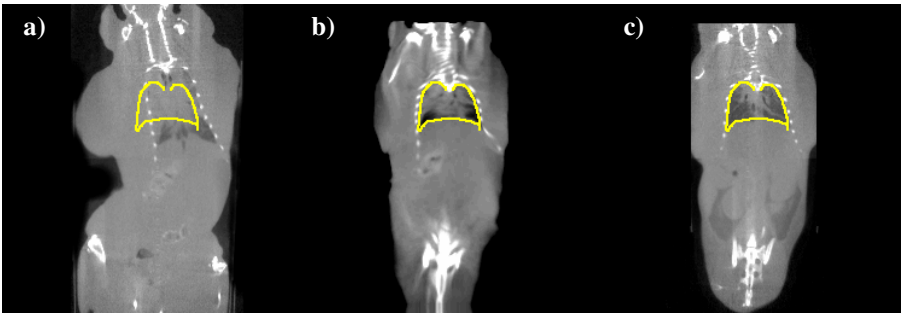



**Fig. 2.** One coronal slice in the source volume (left); the corresponding slice in the target volume (right) , and the transformed source image after registration (middle)

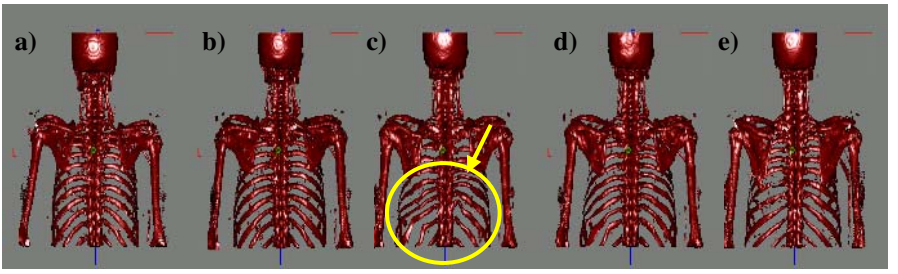



**Fig. 3.** a) Skeleton of the source image, e) skeleton of the target image. b), c), and d) source skeleton registered to target skeleton using a stiff transformation, a very elastic transformation, and two stiffness values, respectively.

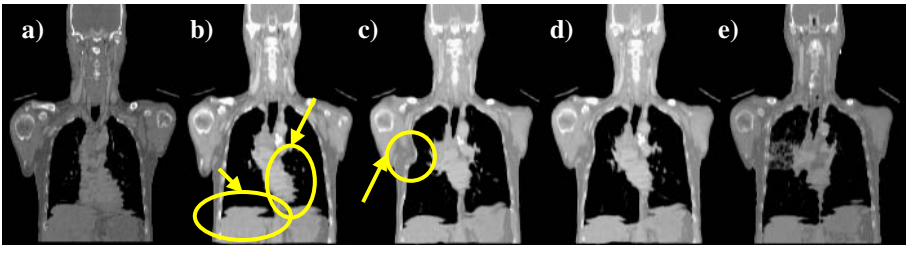



**Fig. 4.** a) One coronal slice in the source volume, e) corresponding slice in the target volume, b), c), and d) source image registered to target image using a stiff transformation, a very elastic transformation, and two stiffness values, respectively

Figure 5 illustrates results we have obtained with another set of upper torso volumes. The left panel shows one sagittal image in one of the volumes (the source). The right panel shows the slice with the same index in the second volume (the target) prior to registration. The second, third, and fourth panels show results obtained with our intensity-based algorithm alone, results obtained with point-based registration alone, and results obtained when both approaches are combined, respectively. The second panel shows typical results obtained when non-rigid registration algorithms cannot be initialized correctly. The overall shape of the registered volume appears correct but

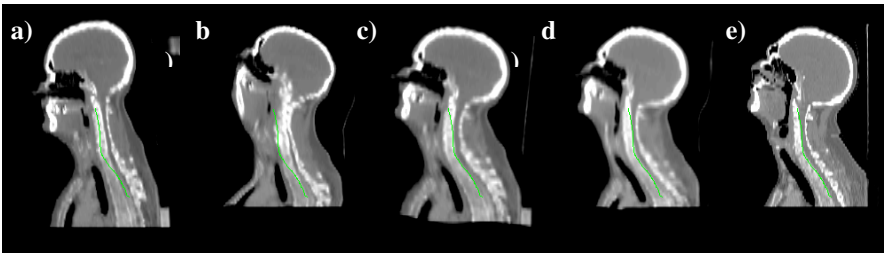



**Fig. 5.** a) One sagittal slice in the source volume, e) the corresponding slice in the target volume, b), c), and d) registration results obtained with intensities alone, points alone, and with both methods combined, respectively

bones have been deformed incorrectly. A closer inspection of the deformation field (not shown here for lack of space) also shows that the deformation field is very irregular. The deformation field obtained with the point-based registration is smooth but the registration relatively inaccurate, as shown in the third panel. As can be seen in this panel, the shape of the head and its size are not exactly similar to those shown in the right panel. Similarly, the sizes of the vertebrae are incorrect. The fourth panel shows that the best results are obtained by combining both approaches.

## 4 Conclusions

In this paper, we present what we believe is the first automatic approach for the registration of articulated structures applicable to inter-subject registration problems. Existing work typically relies on a combination of piecewise rigid body transformations, which requires localizing joints in the image accurately. This is time-consuming and hard to automate. In our method, the process can be fully automated by registering first the entire skeleton using a point-based method that does not require labeling of homologous points. This produces a transformation, which may not be extremely accurate but is nevertheless sufficient to initialize an intensity-based non-rigid registration algorithm. The second step leads to an accurate registration. We also show that better results can be obtained with two stiffness values than with one. Future work includes improving the way points are selected for the point-based registration algorithm and conducting a quantitative evaluation and comparison of these algorithms.

**Acknowledgments.** The small animal CT data sets in this work were provided by Dr. Todd E. Peterson, whose work is supported in part by a Career Award at the Scientific Interface from the BurroughsWellcome Fund. We thank Dr Duggan for providing us the human data sets. This work was supported, in parts, by NIH grant NIBIB 2R01 EB000214-15A1.

## References

1. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," IEEE Transactions on Medical Imaging 16(2), pp. 187-198, 1997.
2. W.M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, “Multi-Modal Volume Registration by Maximization of Mutual Information,” Medical Image Analysis, pp. 35-51, 1996.
3. O Camara, G Delso, I Bloch, “Free Form Deformations Guided by Gradient Vector Flow: A Surface Registration Method in Thoracic and Abdominal PET-CT Applications,” WBIR 2003, LNCS 2717, pp. 224-233, 2003.
4. D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach and D. Hawkes, “Non-rigid registration using free-form deformations: application to breast MR images,” IEEE Trans. Med. Imaging 18, pp. 712–21, 1999.

5. Cai, J., Chu, J.C.H., Recine, D., Sharma, M., Nguyen, C., Rodebaugh, R.,Saxena, V., Ali, A., "CT and PET Lung Image Registration and Fusion in Radiotherapy Treatment Planning Using Chamfer Matching Method," International Journal of Radiation Oncology, Biology, and Physics. 43:883-891, 1999.
6. J. A. Little, D. L. G. Hill, and D. J. Hawkes, "Deformations Incorporating Rigid Structures," Computer Vision and Image Understanding, Vol. 66, No. 2, pp. 223-232, May 1997.
7. Martin-Fernandez MA, Munoz-Moreno E, Martin-Fernandez M, Alberola-Lopez C, "Articulated registration: Elastic registration based on a wire-model," Medical Imaging 2005: Image Processing, SPIE, pp. 182-191, 2005.
8. V. Arsigny, X. Pennec, and N. Ayache, "Polyrigid and Polyaffine Transformations: A New Class of Diffeomorphisms for Locally Rigid or Affine Registration," Proc. of MICCAI'03, pp. 829-837, 2003.
9. X. Papademetris, D.P. Dione, L.W. Dobrucki, L.H. Staib, A.J. Sinusas, "Articulated Rigid Registration for Serial Lower-Limb Mouse Imaging," MICCAI'05, pp. 919-926, 2005.
10. Haili Chui and Anand Rangarajan, "A new point matching algorithm for non-rigid registration," Computer Vision and Image Understanding, pp. 114-141, 2003.
11. S. Gold, A. Rangarajan, C.P. Lu, S. Pappu, E. Mjolsness, "New algorithms for 2D and 3D point matching: pose estimation and correspondence," Pattern Recognition 31(8), pp. 1019-1031, 1998.
12. Gustavo K. Rohde, Akram Aldroubi, and Benoit Dawant, "The adaptive bases algorithm for intensity-based nonrigid image registration," IEEE Trans. Med. Imag., vol. 22, no. 11, pp. 1470-1479, 2003.
13. Valerie Duay, Pierre-François D'Haese, Rui Li, Benoit M. Dawant, "Non-Rigid Registration Algorithm With Spatially Varying Stiffness Properties," ISBI 2004: 408-411.

# Local Intensity Mapping for Hierarchical Non-rigid Registration of Multi-modal Images Using the Cross-Correlation Coefficient

Adrian Andronache, Philippe Cattin, and Gábor Székely

ETH Zurich - Computer Vision Laboratory
Sternwartstrasse 7, CH - 8092 Zurich
{andronache, cattin, szekely}@vision.ee.ethz.ch

**Abstract.** The hierarchical subdivision strategy which decomposes the non-rigid matching problem into numerous local rigid transformations is a very common approach in image registration. For multi-modal images mutual information is the usual choice for the measure of patch similarity. As already recognized in the literature, the statistical consistency of mutual information is drastically reduced when it is estimated for regions covering only a limited number of image samples. This often affects the reliability of the final registration result.

In this paper we present a new intensity mapping algorithm which can locally transform images of different modalities into an intermediate pseudo-modality. Integrated into the hierarchical framework, this intensity mapping uses the local joint intensity histograms of the coarsely registered sub-images and allows the use of the more robust cross-correlation coefficient for the matching of smaller patches.

## 1 Introduction

Medical imaging technologies have become indispensable components in most clinical procedures during the last years. The wide availability of different imaging modalities tremendously increased the need for fast and accurate multi-modal image registration methods. Several surveys and textbooks (e.g. [1, 2, 3, 4, 5] and references therein) have already been published providing a broad and general overview of image registration techniques. The most accurate methods for non-rigid registration are based on physical models but they proved to be computationally very expensive. Therefore, various simplifications have been investigated based on different heuristics to approximate the underlying physical reality by alternative mathematical models. One of these approaches has been proposed by Likar and Pernuš in [6]. They developed a hierarchical image subdivision strategy that decomposes the non-rigid matching problem into an elastic interpolation of numerous local rigid registrations of sub-images of decreasing size. As the local registrations are achieved by maximizing mutual information (MI), the algorithm can be generally applied both for mono- and multi-modal cases. Unfortunately, the usage of MI for image matching has several drawbacks in connection with either interpolation artifacts or the statistical consistency of

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 26–33, 2006.

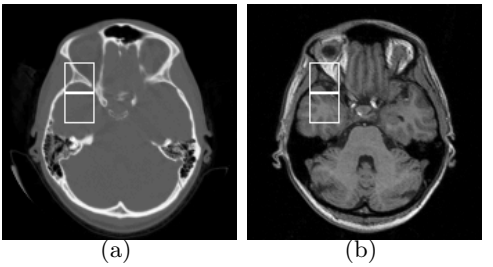
MI (e.g. in [5, 6, 7, 8]). We have demonstrated in [9] that the related problems are becoming increasingly serious during the image subdivision process due to the decreased number of samples used to estimate the two-dimensional joint intensity histogram. This decrease of the MI's statistical consistency weakens the performance of the entire non-rigid registration and limits the number of levels which can be generated during the hierarchical subdivision. It would be therefore desirable to replace MI with a more robust similarity measure. However, the usage of the cross-correlation coefficient (CC) favored by most researchers is restricted to the mono-modal case.

In the past few years, several methods have been proposed either for estimating a functional relationship between the intensities of images from different modalities or for the direct estimation of similarity measures which integrate this functionality in their definition. For example, the VIR criterion presented by Woods in [10] proved to be efficient for matching PET with MR images. In [11] an extension was presented that removed the need for manual segmentation and extended the method's applicability to other modality combinations. Another extension of Woods' VIR criterion called *correlation ratio* is described in [12]. Later on, in [13], an adaptive intensity correction was proposed that combines the correlation ratio with the demons algorithm [14]. A completely different approach for CT-MR cross-registration is described in [15] and bases on a simple intensity mapping of the original CT image such that bone and air have identical appearance as in an MR image. All the proposed methods, however, lead to the appearance of fake structures within the mapped image, which strongly limits their usability. These ghost features caused by imaging details which are not visible in both modalities lead to ambiguities that result in misregistrations.

In this paper we propose a local intensity mapping that allows to switch from MI to the more robust CC at finer levels in the registration hierarchy. In contrast to the already existing approaches that estimate the functional relationship from one image modality to the other, we propose to build a common intermediate *pseudo-modality*. The intensities in both images are mapped simultaneously onto a common contrast space, which is not necessarily one of the two source intensities, but rather a combination of them. Although the transformed images may locally resemble one of the modalities, on an overall scale this is not true. In this paper we present a novel mapping, which is demonstrated on CT/MR image registration but is generally applicable for any combination of modalities.

## 2 Method

The mapping we propose in this paper is relying on the observation that the performance of a registration algorithm will not increase if one of the images contains more structural details than the other. On the contrary, details visible in only one of the images can lead to ambiguities by inducing misleading optima in the similarity measure. The performance of the registration procedure thus only depends on those image features which exist simultaneously in both modalities. The proposed mapping procedure builds an intermediate pseudo-modality of



(a) (b)

**Fig. 1.** Transversal slices of rigidly registered (a) CT and (b) MR acquisitions of the head

the images that will show only the common image features and drop additional details prominent in only one of the modalities.

Figure 1 depicts two corresponding transversal slices from rigidly registered 3D volumes of $512 \times 512 \times 50$ voxels of size $0.39 \times 0.39 \times 0.6$ mm$^3$. Obvious differences can be noticed not only in the intensities of most of the structures but also in the visibility of details of the tissues. The mapping between the intensities is neither linear, nor invertible.

We propose to estimate the functional relationship between the intensities of the different modalities by using the information contained in the joint histogram of the coarsely registered images. In a first step the mean values and the variances of all MR image ($B$) voxels are calculated, which correspond to the normalized intensity values $a = 0..255$ in the CT image ($A$) using the joint histogram $H_{AB}$:

$$\mu_A(a) = \frac{\sum_{b \in B} H_{AB}(a,b) \cdot b}{\sum_{b \in B} H_{AB}(a,b)} \tag{1}$$

$$\sigma_A^2(a) = \frac{\sum_{b \in B} H_{AB}(a,b) \cdot (b - \mu_A(a))^2}{\sum_{b \in B} H_{AB}(a,b)} \tag{2}$$

and likewise for the MR image voxels $b = 0..255$:

$$\mu_B(b) = \frac{\sum_{a \in A} H_{AB}(a,b) \cdot a}{\sum_{a \in A} H_{AB}(a,b)} \tag{3}$$

$$\sigma_B^2(b) = \frac{\sum_{a \in A} H_{AB}(a,b) \cdot (a - \mu_B(b))^2}{\sum_{a \in A} H_{AB}(a,b)}. \tag{4}$$

For each of the histogram bins of the CT and MR, a flag $f_a, f_b$ and a counter $c_a, c_b$ is defined. The flag encodes whether an intensity value should be mapped by the corresponding $\mu$ function or should be kept unchanged:

$$f_a, f_b = \begin{cases} 1 & \text{map to the other modality} \\ 0 & \text{undefined in the joint histogram} \\ -1 & \text{keep the value unchanged.} \end{cases} \tag{5}$$

The counters are increased by one, whenever an intensity from one modality is mapped onto its bin. As it is our aim to suppress structures in a patch which are not visible in the other modality, the flags select intensities according to the image with smaller variance. The counters are also updated according to this decision:

$$\forall a,\ \begin{cases} \text{if } \sigma_A(a) < \sigma_B(\mu_A(a)) \rightarrow f_a = 1, & \text{inc. counter } c_{\mu_A(a)} = c_{\mu_A(a)} + 1 \\ \text{if } \sigma_A(a) > \sigma_B(\mu_A(a)) \rightarrow f_a = -1, & \text{inc. counter } c_a = c_a + 1 \end{cases} \tag{6}$$

and likewise for the MR image:

$$\forall b,\ \begin{cases} \text{if } \sigma_B(b) < \sigma_A(\mu_B(b)) \rightarrow f_b = 1, & \text{inc. counter } c_{\mu_B(b)} = c_{\mu_B(b)} + 1 \\ \text{if } \sigma_B(b) > \sigma_A(\mu_B(b)) \rightarrow f_b = -1, & \text{inc. counter } c_b = c_b + 1. \end{cases} \tag{7}$$

As can be seen in the schematic joint histogram in Fig. 2(a) three different regions can be distinguished according their variance. For the regions where $f_a > f_b$ and $f_a < f_b$ the mapping direction is unambiguous and indicated with an arrow. In Fig. 2(b) the same regions are labeled according to the flag notation. If $\sigma_A \approx \sigma_B$ no clear decision can be made. For the intensities in this ambiguous region (see Fig. 2(c)) it is very likely that the value $b_l$ will be mapped to $a_i$ and $a_i$ which in turn is associated with $b_k$. Accordingly, the mapping function $\nu_A(a)$ for $A$ can be written as:

$$\forall a = 0..255,\ b = \nu_A(a) = \begin{cases} \text{if } f_a > f_{\mu_A(a)} \rightarrow \mu_A(a) \\ \text{if } f_a < f_{\mu_A(a)} \rightarrow a \\ \text{if } f_a = f_{\mu_A(a)} \rightarrow \text{ambiguity.} \end{cases} \tag{8}$$

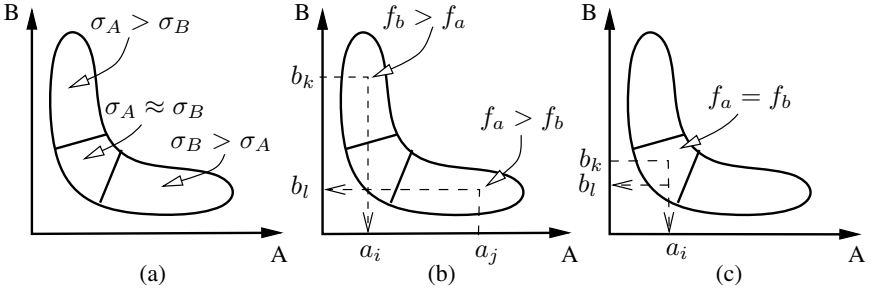



**Fig. 2.** Schematic joint histogram with (a) three regions defined by their variances, (b) the three regions labeled with the flag notation, and (c) illustration of ambiguities

Two different ambiguous cases can be distinguished: (1) $f_a = f_{\mu_A(a)} = 1$, i.e. the intensities in both modalities are to be changed and (2) $f_a = f_{\mu_A(a)} = -1$ i.e. both of them should be kept. The mapping counters $c_a$, $c_{\mu_A(a)}$ are used to resolve such situations:

$$\forall f_a = f_{\mu_A(a)},\ b = \nu_A(a) = \begin{cases} \text{if } c_a > c_{\mu_A(a)} \rightarrow a \\ \text{if } c_a < c_{\mu_A(a)} \rightarrow \mu_A(a). \end{cases} \tag{9}$$

Corresponding formulas for Eq. 8/9 are used for $\nu_B(b)$. The ambiguities have to be resolved iteratively until no further changes in the resulting mappings are detected. Figure 3 and 4 show examples of CT/MR mappings using the proposed method.

## 3 Results

Two representative examples were chosen to demonstrate the advantages of integrating this intensity mapping procedure into the hierarchical registration, such that CC can be used as the similarity measure instead of MI after a certain level of the hierarchy has been reached. Two regions of interest have been selected for illustration, marked with white squares on Fig. 1. All patches are of $64 \times 64 \times 17$ voxels, equivalent to the $4^{\text{th}}$ level of the subdivision.

The first experiment was performed with an image pair (upper white squares in Fig. 1) containing rich structural details. Figure 3 shows the original patches, their intensity mapped versions and the behavior of MI (on the original) and CC (on the intensity mapped images) for horizontal displacements up to $\pm 10$ pixels. It can be seen that for regions having sufficient structural information, both similarity measures are sufficiently stable for finding the correct registration position.

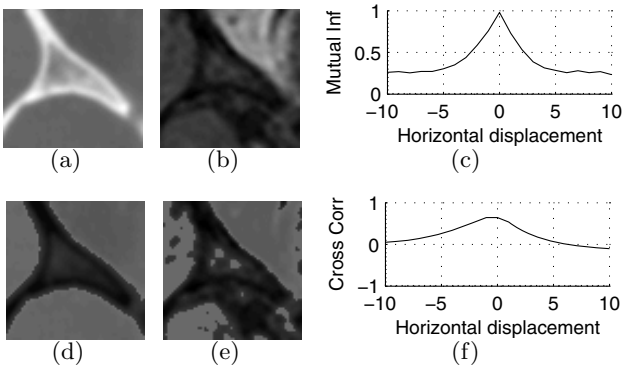



**Fig. 3.** (a,b) Initial patches showing rich structural details and (d,e) their intensity mapped versions. (c) The response of MI on the original and (f) CC on the intensity mapped images to horizontal translations.

A different region of interest (lower white squares in Fig. 1) has been used for the same experiment. While the corresponding CT patch is almost uniform, the MR image shows significant contrast within the brain tissue covered. This is a classical case in which MI generally fails to find the correct registration position, see [9]. Figure 4 shows the original and intensity mapped patches together with the comparison between the MI and the CC responses to horizontal translations. While CC remains robust for this region, too, MI shows highly unreliable behavior. Note, that our hierarchical strategy partitions only the floating image and

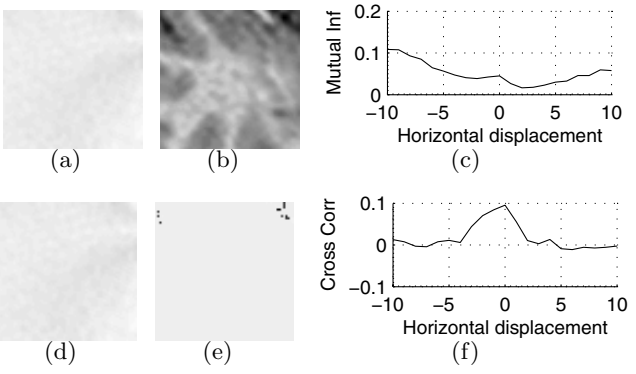



**Fig. 4.** (a,b) Patches with major differences of tissue contrast in CT and MR. (d,e) Intensity mapped versions of the images. (c) The response of MI on the original and (f) CC on the intensity mapped images to horizontal translations.

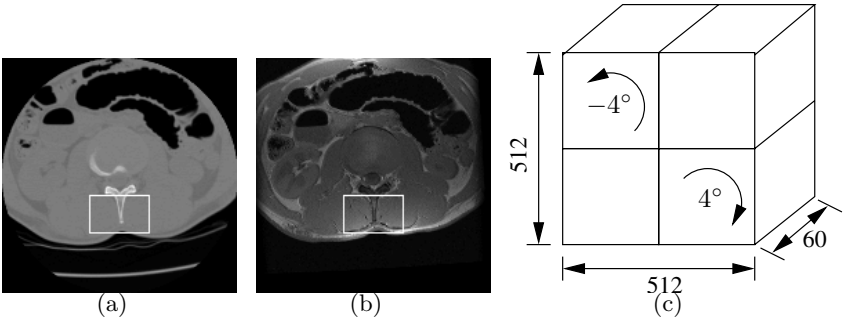



**Fig. 5.** (a) CT and (b) MR sample slice of the spine volume used for the validation tests (c) schematic of the artificial deformation field

the local similarity measure for a partitioned sub-image is calculated from its volume of overlap with the entire reference image. Therefore, only the sub-images on the border of the volume are effected by an eventual change in the overlapping volume. According to our experience, this leads seldom to misregistrations which can fully be corrected by the subsequent regularization step.

In order to quantitatively analyze the advantages of integrating the proposed intensity mapping strategy into the hierarchical non-rigid registration procedure, an artificial registration scenario which consists of recovering a predefined deformation field was used. The underlying pre-registered CT/MR datasets of the spine had a size of $512 \times 512 \times 60$ voxels of dimension $0.47 \times 0.47 \times 1.25$ mm$^3$, see Fig. 5(a,b). The CT dataset was split in 4 equally sized blocks of size $256 \times 256 \times 60$ voxels. Two of these blocks were rigidly rotated as shown in Fig. 5(c). The deformed volume was then interpolated using thin plate splines (TPS) and partial volumes. The MR volume was then registered using (1) MI

**Table 1.** Registration error calculated for the entire volume as well as for the local region marked with a white square in Fig. 5(a,b)

| Results | Direction | Similarity Measure MI | MI+CC | Improvement |
|---|---|---|---|---|
| entire volume | in plane ($xy$) | $0.5157 \pm 0.4137$ mm | $0.4795 \pm 0.4617$ mm | 7.02% |
| | out of plane ($z$) | $0.4262 \pm 0.5615$ mm | $0.3990 \pm 0.4423$ mm | 6.38% |
| | overall ($xyz$) | $0.7380 \pm 0.6240$ mm | $0.6790 \pm 0.5804$ mm | 7.99% |
| local region | in plane ($xy$) | $0.3641 \pm 0.2205$ mm | $0.2377 \pm 0.1782$ mm | 34.71% |
| | out of plane ($z$) | $0.2768 \pm 0.2743$ mm | $0.2379 \pm 0.2085$ mm | 14.05% |
| | overall ($xyz$) | $0.4987 \pm 0.2904$ mm | $0.3653 \pm 0.2342$ mm | 26.74% |

during the entire hierarchical registration procedure and (2) when switching MI to CC at the $4^{\text{th}}$ hierarchical level. The recovered deformation fields were then compared to the known artificial deformation field.

Table 1 summarizes the average and standard deviation of the registration error for the entire dataset and for the region marked with a white square in Fig. 5(a,b). As the gain of switching from MI to CC only applies to a small number of sub-images, the average registration error over the entire volume improves only slightly. However, the registration error improved up to 34.71% for the selected region ($80 \times 120 \times 40$ voxels) where MI generally tends to fail.

## 4 Conclusions

As has been previously discussed in the literature, MI shows unsatisfactory behavior for the matching of structureless or small image patches due to the lack of statistical consistency caused by the small number of available image samples. CC proved to be more robust, but it can not be directly used for cases, where the intensity relation between the modalities is non-linear. The mapping strategy presented by this paper enables the combination of both similarity measures for multi-modal registration procedures relying on a hierarchical subdivision strategy. At the first levels of the hierarchy, where the partitions are still relatively large, MI can be used to coarsely register the corresponding patches. After this stage, the images can be transformed to a pseudo-modality using the presented mapping technique and the similarity measure can be switched to the more robust CC.

With the proposed hybrid approach that uses MI for the first levels and CC for the last few levels, two important properties of these similarity measures can be seamlessly combined in a unique manner, namely the multi-modal capabilities of MI with the robustness of CC without increasing the computational complexity of the underlying algorithm.

## Acknowledgments

This work has been supported by the CO-ME/NCCR research network of the Swiss National Science Foundation (http://co-me.ch).

## References


1. Maintz, J., Viergever, M.: A survey of medical image registration. Medical Image Analysis **2** (1998) 1–36
2. Hajnal, J.V., Hill, D.L., Hawkes, D.J.: Medical Image Registration. CRC Press (2001)
3. Hill, D.L.G., Batchelor, P.G., Holden, M., Hawkes, D.J.: Medical image registration. PHYSICS IN MEDICINE AND BIOLOGY **46** (2001)
4. Maes, F., Vandermeulen, D., Suetens, P.: Medical image registration using mutual information. Proc IEEE **91** (2003) 1699–1722
5. Pluim, J.P.W., Maintz, J.A., Viergever, M.A.: Mutual-information-based registration of medical images: a survey. IEEE Transactions on Medical Imaging **22** (2003) 986–1004
6. Likar, B., Pernuš, F.: A hierarchical approach to elastic registration based on mutual information. Image and Vision Computing **19** (2001) 33–44
7. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging **16** (1997) 187–198
8. Pluim, J.P.W., Maintz, J.A., Viergever, M.A.: Mutual information matching and interpolation artefacts. Proc SPIE Medical Imaging **3661** (1999) 56–65
9. Andronache, A., Cattin, P., Székely, G.: Adaptive subdivision for hierarchical non-rigid registration of multi-modal images using mutual information. In Duncan, J.S., Gerig, G., eds.: MICCAI (2). Volume 3750., Springer-Verlag GmbH (2005) 976–983
10. Woods, R.P., Mazziotta, J.C., Cherry, S.R.: MRI-PET registration with automated algorithm. Comput. Assist. Tomogr. **17** (1993) 536–546
11. Nikou, C., Heitz, F., Armspach, J., , Namer, I.: Single and multimodal subvoxel registration of dissimilar medical images using robust similarity measures. SPIE Conference on Medical Imaging (1998)
12. Roche, A., Malandain, G., Pennec, X., Ayache, N.: The correlation ratio as a new similarity measure for multimodal image registration. In et al., W.W., ed.: MICCAI. (1998)
13. Guimond, A., Roche, A., Ayache, N., Meunier, J.: Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. IEEE Transactions on Medical Imaging **20** (2001) 58–69
14. Thirion, J.: Non-rigid matching using demons. In: Int. Conf. Computer Vision and Pattern Recognition. (1996)
15. van den Elsen, P.A., Pol, E.J.D., Sumanaweera, T.S., Hemler, P.F., Napel, S., Adler, J.R.: Grey value correlation techniques for automatic matching of CT and MR brain and spine images. Visualization in Biomed. Comp. **2359** (1994) 227–237

# Multi-modal Image Registration Using Dirichlet-Encoded Prior Information

Lilla Zöllei and William Wells

MIT, CSAIL; 32 Vassar St (Bld 32), Cambridge, MA 02139, USA
{lzollei, sw}@csail.mit.edu

**Abstract.** We present a new objective function for the registration of multi-modal medical images. Our novel similarity metric incorporates both knowledge about the current observations and information gained from previous registration results and combines the relative influence of these two types of information in a principled way. We show that in the absence of prior information, the method reduces approximately to the popular entropy minimization approach of registration and we provide a theoretical comparison of incorporating prior information in our and other currently existing methods. We also demonstrate experimental results on real images.

## 1 Problem Definition

Multi-modal intensity-based registration of medical images can be a challenging task due to great differences in image quality and resolution of the input data sets. Therefore, besides using the intensity values associated with the current observations, there have been attempts using certain statistics established at the correct alignment of previous observations in order to increase the robustness of the algorithms [4, 1]. One such example is applying the joint probability of previously registered images as a model for the new set of images. That approach requires one to assume that we have access to the joint distributions of previously registered images and also that the resulting joint distribution model accurately captures the statistical properties of other unseen image pairs at registration. The accuracy of such methods, however, is biased by the quality of the model. This motivates an approach which is both model-reliant and model-free in order to guarantee both robustness and high alignment accuracy. Such ideas have been recently formulated by Chung[1] and Guetter [3]. Chung et al. have proposed the sequential utilization of a Kullback Leibler (KL)-divergence and a Mutual Information (MI) term, while Guetter et al. incorporate the same two metrics into a simultaneous optimization framework. In both methods there is an arbitrary parameter that decides how to balance the two influences.

[1] Private communications with Prof A. Chung (The Hong Kong University of Science and Technology).

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 34–42, 2006.

## 2 Proposed Method

We formulate the registration task by balancing the contributions of data and prior terms in a principled way. We follow a Bayesian framework and introduce a prior on joint probability models reflecting our confidence in the quality of these statistics learned from previously registered image pairs.

We define a normalized likelihood-based objective function on input data sets $u$ and $v$ by optimizing over *both* transformation ($T$) and the parameters of the unknown joint probability model ($\Theta$):

$$\arg\max_{T,\Theta} \mathcal{F}(T,\Theta) = \arg\max_{T,\Theta} \frac{1}{N} \log p([u, v_T]; T, \Theta). \tag{1}$$

If we assume that $\Theta$ encodes information about intensity value joint occurrences as parameters of an unknown multinomial model in $N$ independent trials and let the random vector $Z = \{Z_1, ..., Z_g\}$ indicate how many times each event (joint occurrence of corresponding intensity values) occurs, then $\sum_{i=1}^{g} Z_i = N$, where $N$ is the size of the overlapping region of the observed images. Also, the probability distribution of the random vector $Z \sim \text{Multinom}(\text{N}; \Theta)$ is given by

$$P(Z_1 = z_1, ..., Z_g = z_g) = \frac{N!}{\prod_{i=1}^{g} z_i!} \prod_{i=1}^{g} \theta_i^{z_i}. \tag{2}$$

According to this interpretation, $Z$ summarizes the event space of the joint intensity samples $[u, v_T]$ and $N$ indicates the observed sample size. Such a representation is convenient as the $\theta_i$ parameters naturally correspond to the parameters of the widely used histogram encoding of the joint statistics of images. Given $g$ number of bins, the normalized contents of the histogram bins are $\Theta = \{\theta_1, ..., \theta_g\}$ with $\theta_i \geq 0$ and $\sum_{i=1}^{g} \theta_i = 1$. Additionally, given the multinomial representation, prior information about the bin contents can be expressed by using Dirichlet distribution, the conjugate prior to a multinomial distribution.

Dirichlet distributions are multi-parameter generalizations of the Beta distribution. They define a distribution over distributions, thus the result of sampling a Dirichlet is a multinomial distribution in some discrete space. In the case where $\Theta = \{\theta_1, ..., \theta_g\}$ represents a probability distribution on the discrete space, the Dirichlet distribution over $\Theta$ is often written as

$$\text{Dirichlet}(\Theta; w) = \frac{1}{Z(w)} \prod_{i=1}^{g} \theta_i^{(w_i - 1)} \tag{3}$$

where $w = \{w_1, w_2, ..., w_g\}$ are the Dirichlet parameters and $\forall w_i > 0$. The normalization term is defined as

$$Z(w) = \frac{\prod_{i=1}^{g} \Gamma(w_i)}{\Gamma(\sum_{i=1}^{g} w_i)}, \tag{4}$$

where

$$\Gamma(w_i) = \int_0^{\infty} t^{w_i - 1} e^{-t} dt. \tag{5}$$

We, however, use another encoding of the distribution. We assign $w_i = \alpha m_i$, where $\alpha > 0$ and $\sum_{i=1}^{g} m_i = 1$. Accordingly,

$$\text{Dirichlet}(\Theta; \alpha, M) = \frac{1}{Z(\alpha M)} \prod_{i=1}^{g} \theta_i^{(\alpha m_i - 1)} = \frac{\Gamma(\alpha)}{\prod_{i=1}^{g} \Gamma(\alpha m_i)} \prod_{i=1}^{g} \theta_i^{(\alpha m_i - 1)}. \quad (6)$$

This representation is more intuitive, as we can interpret $M = \{m_1, m_2, ..., m_g\}$ as a set of *base measures* which, it turns out, are also the mean value of $\Theta$, and $\alpha$ is a *precision* parameter showing how concentrated the distribution around $M$ is. We can also think of $\alpha$ as the number of pseudo measurements observed to obtain $M$. The higher the former number is, the greater our confidence becomes in the values of $M$. When using a Dirichlet distribution, the expected value and the variance of the $\Theta$ parameters can be obtained in closed form [2]. They are

$$\text{E}(\theta_i) = m_i \quad \text{and} \quad \text{Var}(\theta_i) = \frac{m_i(1 - m_i)}{\alpha(\alpha + 1)}. \quad (7)$$

Later we also need to compute the logarithm of this distribution which is

$$\log\left[\text{Dirichlet}(\Theta; \alpha, M)\right] = \log\left[\frac{1}{Z(\alpha M)} \prod_{i=1}^{g} \theta_i^{(\alpha m_i - 1)}\right] \quad (8)$$

$$= \log \prod_{i=1}^{g} \theta_i^{(\alpha m_i - 1)} - \log\left[Z(\alpha M)\right] \quad (9)$$

$$= \sum_{i=1}^{g} \log(\theta_i^{(\alpha m_i - 1)}) - \log\left[Z(\alpha M)\right] \quad (10)$$

$$= \sum_{i=1}^{g} (\alpha m_i - 1) \log \theta_i - \log\left[Z(\alpha M)\right]. \quad (11)$$

Thus incorporating the prior model we can write a new objective function as:

$$\arg\max_{T,\Theta} \mathcal{F}(T, \Theta) =$$

$$= \arg\max_{T,\Theta} \frac{1}{N} \log[p([u, v_T]; T|\Theta)p(\Theta)] \quad (12)$$

$$= \arg\max_{T,\Theta} \frac{1}{N} \left[\log p([u, v_T]; T|\Theta) + \log \text{Dirichlet}(\Theta; \alpha M)\right] \quad (13)$$

$$= \arg\max_{T,\Theta} \frac{1}{N} \left[\log p([u, v_T]; T|\Theta) + \sum_{i=1}^{g} (\alpha m_i - 1) \log \theta_i - \log Z(\alpha M)\right]. \quad (14)$$

We may choose to order the optimization of $T$ and $\Theta$ and require that only the optimal transformation $T$ be returned. We denote the distribution parameters that maximize the expression in brackets to be optimized as $\hat{\Theta}_T$. This, in fact, corresponds to the MAP parameter estimate of the multinomial parameters given the image data and some value of T. If we indicate the KL divergence of

distributions as $D$ and the Shannon entropy measure as $H$, the aligning transformation $T_{\text{DIR}}$ can be expressed as [12]:

$$T_{\text{DIR}} = \arg\max_{T} \left[ \frac{1}{N} \sum_{i=1}^{N} \log p(u(x_i), v(T(x_i)); T|\hat{\Theta}_T) + \frac{1}{N} \sum_{i=1}^{g} (\alpha m_i - 1) \log \hat{\theta}_{T_i} \right] \tag{15}$$

$$\approx \arg\min_{T} \left[ D(p_T || \hat{p}_{T,\hat{\Theta}_T}) + H(p_T) - \frac{1}{N} \sum_{i=1}^{g} (\alpha m_i - 1) \log \hat{\theta}_{T_i} \right], \tag{16}$$

where $p_T$ is the true probability distribution of the input observations given parameter $T$ and $\hat{p}_{T,\hat{\Theta}_T}$ is the estimated model joint distribution parameterized by $T$ and $\hat{\Theta}_T$. The newly proposed objective function can be interpreted as the composition of a data- and a prior-related term. The former expresses discrepancies between the true source distribution and its estimated value, while the latter incorporates knowledge from previous correct alignments. As it might not be intuitive how that information influences the alignment criterion, in the following, we further manipulate the third term in Eq.(16). The prior-related term in Eq.(16) can be expanded into a sum of two terms:

$$-\frac{1}{N} \sum_{i=1}^{g} (\alpha m_i - 1) \log \hat{\theta}_{T_i} = -\frac{\alpha}{N} \sum_{i=1}^{g} m_i \log \hat{\theta}_{T_i} + \frac{1}{N} \sum_{i=1}^{g} \log \hat{\theta}_{T_i}. \tag{17}$$

If we assume that both the base parameters of the Dirichlet distribution $M = \{m_1, ..., m_g\}$ and the $\Theta = \{\theta_1, ..., \theta_g\}$ parameters represent normalized bin contents of histogram encodings of categorical probability distributions $\mathcal{P}_M$ and $\mathcal{P}_{\hat{\Theta}_T}$, respectively, and furthermore, if we denote a uniform categorical probability distribution function by $\mathcal{P}_U$ where each of the $g$ number of possible outcomes equals $\left(\frac{1}{g}\right)$, then we can approximate the prior-related term through:

$$-\frac{\alpha}{N} \sum_{i=1}^{g} m_i \log \hat{\theta}_{T_i} + \frac{1}{N} \sum_{i=1}^{g} \log \hat{\theta}_{T_i} =$$

$$= \frac{\alpha}{N} \left[ D(\mathcal{P}_M \| \mathcal{P}_{\hat{\Theta}_T}) + H(\mathcal{P}_M) \right] + \frac{1}{N} \sum_{i=1}^{g} \log \hat{\theta}_{T_i} \tag{18}$$

$$= \frac{\alpha}{N} \left[ D(\mathcal{P}_M \| \mathcal{P}_{\hat{\Theta}_T}) + H(\mathcal{P}_M) \right] - \frac{g}{N} \left[ D(\mathcal{P}_U \| \mathcal{P}_{\hat{\Theta}_T}) + H(\mathcal{P}_U) \right]. \tag{19}$$

After dropping terms that are constant over the optimization, the objective function from Eq.(16) can then be expressed as

$$T_{\text{DIR}} \approx \arg\min_{T} \left[ D(p_T \| \hat{p}_{T,\hat{\Theta}_T}) + H(p_T) + \frac{\alpha}{N} D(\mathcal{P}_M \| \mathcal{P}_{\hat{\Theta}_T}) - \frac{g}{N} D(\mathcal{P}_U \| \mathcal{P}_{\hat{\Theta}_T}) \right]. \tag{20}$$

Therefore, our new registration objective can be interpreted as the weighted sum of four information theoretic terms. We refer to them as the *data* terms,

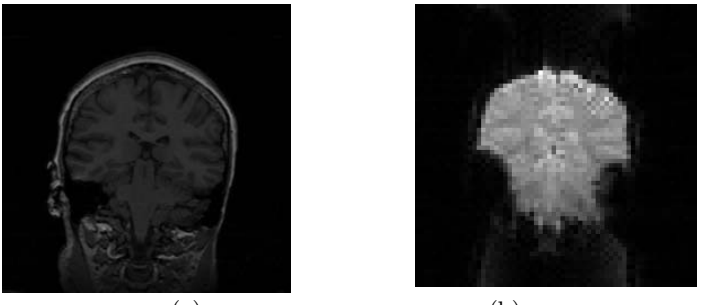
the *prior* term and the estimation term. The first two terms, the data-related terms, indicate how well the observations fit the model given optimal distribution parameters $\hat{\Theta}_T$. The third term measures the KL-divergence between two categorical distributions over the parameters describing the *pseudo* and the current observations and the fourth term evaluates the KL-divergence between two other categorical distributions, the uniform and the one characterizing the parameters of the current observations. Note, as the uniform distribution has the highest entropy among all, maximizing the KL-divergence from it is very similar to minimizing the entropy of the distribution. As $N$ is fixed and given by the number of the observed input intensity pairs, the weighting proportion depends solely on $\alpha$, the *precision* parameter of the Dirichlet distribution. It is this value that determines how much weight is assigned to the *prior* term or in other words it ensures that the mode of the prior is centered on the previously observed statistics. That arrangement is intuitively reasonable: when $\alpha$ is high, the Dirichlet base counts are considered to originate from a large pool of previously observed, correctly aligned data sets and thus we have high confidence in the prior; when $\alpha$ is low, prior observations of correct alignment are restricted to a smaller number of data sets thus the prior is trusted to a lesser extent. Interestingly, most often when one relies on fixed model densities, it is exactly this $\alpha$ value that is missing, i.e. there is no notion about how many prior registered data sets have been observed in order to construct the known model distribution. We also point out that by discarding the prior information and assuming that the distribution estimation process is sufficiently accurate, the objective function is approximately equivalent to the joint entropy registration criterion.

## 3 Preliminary Probing Experiments

In order to experimentally verify the previously claimed advantages of our novel registration algorithm, we designed a set of probing experiments to describe the capture range and accuracy of a set of objective functions. A probing experiment corresponds to the detailed characterization of an objective function with respect to certain transformation parameters. It helps to describe the capture range (the interval over which the objective function does not contain any local optima besides the solution) and accuracy of the objective function.

We compared the behavior of our method to that of three others: joint entropy [9], negative mutual information [5, 11] and KL-divergence [1]. The first two of these methods only consider data information and the third one relies on previous registration results. Ours incorporates both. It thus benefits from the current observations allowing for a well-distinguishable local extrema at correct alignment and also from previous observations increasing the capture range of the algorithm.

The input data sets were 2D acquisitions of a Magnetic Resonance Imaging (MRI) and an echoplanar MRI (EPI) image (see Fig. 1). Historically, the registration of these two modalities has been very challenging because of the low contrast information in the latter [8]. We carried out the probing experiments in

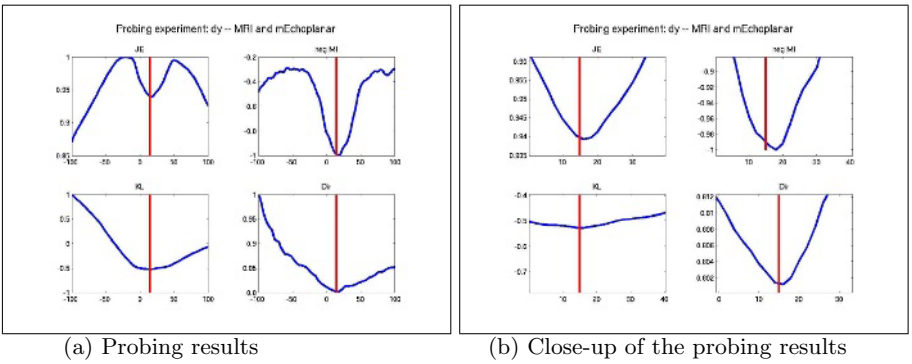


(a) (b)

**Fig. 1.** 2D slices of a corresponding (a) MRI and (b) EPI data set pair

the y- (or vertical) direction. This is the parameter along which a strong local optimum occurs in the case of all the previously introduced objective functions. In order to avoid any biases towards the zero solution, we offset the input EPI image by 15 mm along the probing direction. Thus the local optimum is expected to be located at this offset position – and not at zero – on the probing curves. The objective functions were all evaluated in the offset interval of [-100, 100] mm given 1 mm step sizes.

The probing experiment results are displayed in Fig. 2. In the case of joint entropy (JE), we find a close and precise local optimum corresponding to the offset solution location. However, the capture range is not particularly wide; beyond a narrow range of offset, several local optima occur. In the case of negative MI, the capture range is just a bit wider. The KL objective function, as expected, increases the capture range. Nevertheless, its accuracy in locating the offset optimal solution is not sufficient. In fact, around the expected local minimum the curve of the objective function is flat thus preventing the precise localization of




(a) Probing results (b) Close-up of the probing results

**Fig. 2.** Probing results related to four different objective functions: joint entropy, MI, KL, our method (top-to-bottom, left-to-right)

the solution. The probing curve of our novel similarity metric demonstrates both large capture range and great accuracy. Thus relying on both previous registration results and the current observations, this new metric is able to eliminate the undesired local minimum solutions.

## 4 Connecting the Dirichlet Encoding to Other Prior Models

Finally, we diverge slightly from our main analysis. We draw similarities between the Dirichlet and other encodings of prior information on distribution parameters. Such an analysis facilitates a better understanding of the advantages of the Dirichlet encoding and it creates a tight link with other methods.

We start our analysis by showing that the maximum likelihood solution for the multinomial parameters $\Theta$ is equivalent to the histogrammed version of the observed intensity pairs drawn from the corresponding input images. Then, using these results, we demonstrate that the MAP estimate of the multinomial parameters (with a Dirichlet prior on them) is the histogram of the pooled data, which is the combination of the currently observed samples and the hypothetical prior counts encoded by the Dirichlet distribution.

### 4.1 ML Solution for Multinomial Parameters

In this section, we rely on the relationship that the joint distribution of the observed samples can be obtained by joint histogramming and the normalized histogram bin contents can be related to the parameters of a multinomial distribution over the random vector $Z$. Then the probability distribution of the random vector $Z \sim \text{Multinom}(\text{N}; \Theta)$ is given by

$$P(Z_1 = z_1, ..., Z_g = z_g) = \frac{N!}{\prod_{i=1}^{g} z_i!} \prod_{i=1}^{g} \theta_i^{z_i}. \tag{21}$$

Again, according to this interpretation, $Z$ summarizes the event space of the joint intensity samples $[u, v_T]$ and $N$ indicates the observed sample size. If we want to then optimize the log version of this expression with respect to the $\Theta$ parameter, we write

$$\hat{\Theta} = \arg\max_{\Theta} \log \frac{N!}{\prod_{i=1}^{g} z_i!} \prod_{i=1}^{g} \theta_i^{z_i} \tag{22}$$

$$= \arg\max_{\Theta} \sum_{i=1}^{g} z_i \log \theta_i. \tag{23}$$

In other words, when searching for the maximum likelihood parameters of multinomial parameters, we need to compute the mode of the expression in Eq. (23) over all $\theta_i$'s. This formulation is very similar to that of the logarithm of the Dirichlet distribution which we formulated in Eq.(11). From probability theory

we know that the mode of that expression is taken at $\left[\frac{\alpha m_i - 1}{\alpha - g}\right]$. Thus if we define $(\alpha m_i \equiv z_i + 1)$, the mode of Eq. (23) is found at

$$\hat{\theta_i} = \frac{\alpha m_i - 1}{\alpha - g} = \frac{(z_i + 1) - 1}{\sum_{i=1}^{g} z_i} = \frac{z_i}{\sum_{i=1}^{g} z_i}. \tag{24}$$

Accordingly, the optimal $\theta_i$ parameter – in the maximum likelihood sense – is the one that can be computed by the number of corresponding counts normalized by the total number of counts. That is exactly the approximation that is utilized by the popular histogramming approach. Therefore, we can state that the maximum likelihood solution for the multinomial parameters is achieved by histogramming.

### 4.2 MAP Solution for Multinomial Parameters with a Dirichlet Prior

In this section we return to the MAP problem formulation that originated our analysis. Here, in order to find the optimal set of distribution parameters $\hat{\Theta}$ with a prior assigned to them, we have

$$\hat{\Theta} = \arg\max_{\Theta} \sum_{i=1}^{g} z_i \log\theta_i + \sum_{i=1}^{g} (\alpha m_i - 1) \log\theta_i \tag{25}$$

$$= \arg\max_{\Theta} \sum_{i=1}^{g} (z_i + \alpha m_i - 1) \log\theta_i \tag{26}$$

If we now define $\alpha' m'_i \equiv z_i + \alpha m_i$, then the above simplifies to

$$\hat{\theta_i} = \frac{\alpha' m'_i - 1}{\sum_{i=1}^{g} (\alpha' m'_i) - k} = \frac{z_i + \alpha_i - 1}{\sum_{i=1}^{g} (\alpha m_i) - k} = \frac{z_i + c_i}{\sum_{i=1}^{g} z_i + c_i}. \tag{27}$$

where $c_i = (\alpha m_i - 1)$ are *counting* parameters related to the *pseudo* counts of the Dirichlet distribution. That is to say, the optimal $\theta_i$ parameter – in the maximum a posteriori sense – is the one that can be computed by the sum of the corresponding observed and pseudo counts normalized by the total number of observed and pseudo counts. In other words, in order to compute the optimal $\theta_i$ parameter, we need to pool together the actually observed and the pseudo counts and do histogramming on this merged collection of data samples.

Interestingly enough, this formulation forms a close relationship with another type of entropy-based registration algorithm. Sabuncu et al. introduced a registration technique based upon minimizing Renyi entropy, where the entropy measure is computed via a non-plug-in entropy estimator [7, 6]. This estimator is based upon constructing the EMST (Euclidean Minimum Spanning Tree) and using the edge length in that tree to approximate the entropy. According to their formulation, prior information is introduced into the framework by *pooling* together corresponding samples from the aligned (prior distribution model) and from the unaligned (to be registered) cases. Throughout the optimization, the

model observations remain fixed and act as anchor points to bring the other samples into a more likely configuration. The reason why such an arrangement would provide a favorable solution has not been theoretically justified. Our formulation gives a proof for why such a method strives for the optimal solution.

Very recently, another account of relying on pooling of prior and current observations been published [10]. The authors use this technique to solve an MRI-CT multi-modal registration task.

## Acknowledgement

This work was supported by NIH grants 5 P41 RR13218, U54 EB005149, and 1U41RR019703-01A2, and by NSF grant EEC 9731748.

## References

1. A.C.S. Chung, W.M.W. Wells III, A. Norbash, and W.E.L. Grimson. Multi-modal Image Registration by Minimizing Kullback-Leibler Distance. In *MICCAI*, volume 2 of *LNCS*, pages 525–532. Springer, 2002.
2. A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis*. CRC Press LLC, Boca Raton, FL, 2003.
3. C. Gütter, C. Xu, F. Sauer, and J. Hornegger. Learning based non-rigid multimodal image registration using kullback-leibler divergence. In G. Gerig and J. Duncan, editors, *MICCAI*, volume 2 of *LNCS*, pages 255–263. Springer, 2005.
4. M. Leventon and W.E.L. Grimson. Multi-modal Volume Registration Using Joint Intensity Distributions. In *MICCAI*, LNCS, pages 1057–1066. Springer, 1998.
5. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.
6. M.R. Sabuncu and P.J. Ramadge. Gradient based optimization of an emst registration function. In *Proceedings of IEEE Conference on Acoustics, Speech and Signal Processing*, volume 2, pages 253–256, March 2005.
7. M.R. Sabuncu and P.J. Ramadge. Graph theoretic image registration using prior examples. In *Proceedings of European Signal Processing Conference*, September 2005.
8. S. Soman, A.C.S. Chung, W.E.L. Grimson, and W. Wells. Rigid registration of echoplanar and conventional magnetic resonance images by minimizing the kullback-leibler distance. In *WBIR*, pages 181–190, 2003.
9. C. Studholme, D.L.G. Hill, and D.J. Hawkes. An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, 32(1):71–86, 1999.
10. M. Toews, D. L. Collins, and T. Arbel. Maximum a posteriori local histogram estimation for image registration. In *MICCAI*, volume 2 of *LNCS*, pages 163–170. Springer, 2005.
11. W.M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1:35–52, 1996.
12. L. Zöllei. *A Unified Information Theoretic Framework for Pair- and Group-wise Registration of Medical Images*. PhD thesis, Massachusetts Institute of Technology, January 2006.

# Removal of Interpolation Induced Artifacts in Similarity Surfaces

Olivier Salvado[1] and David L. Wilson[1,2]

[1] Department of Biomedical Engineering,
Case Western Reserve University, Cleveland, OH 44106
{olivier.salvado, david.wilson}@case.edu
[2] Department of Radiology,
University Hospitals of Cleveland, Cleveland, OH 44106

**Abstract.** Registration of two images requires interpolation to generate a new image on a transformed grid, and the optimal transformation that maps an image to the other is found by maximizing a similarity measure. Similarity surfaces are subject to scalloping artifacts due to interpolation that give local maxima, and, in some cases, erroneous global maxima. We propose a new linear filter that is applied to input images and which removes scalloping artifacts from cross-correlation and mutual-information similarity surfaces. The computational burden is sufficiently low that it can be used in every iteration of an optimization process. In addition, this new filter generates image data with constant variance after linear interpolation, making measurements of signal change more reliable. Following filtering of MR images, similarity surfaces are smoothed with removal of local maxima and biased global maxima.

## 1 Introduction

In intensity based registration, one finds the transformation that maximizes a similarity measure which quantifies the alignment of two data sets (see [1] for a recent survey). Typical examples of similarity measures are mutual information and cross-correlation [2]. Significant scalloping artifacts can be observed in the cross-correlation and mutual-information surfaces. Two major problems arise from these perturbations: local maxima may hinder any optimization algorithm, and global maximum may happen not for the optimal transformation parameters.

Some studies have determined the origins of perturbations. During registration, transformation of the floating image to the reference image requires an interpolation step and often linear interpolation is used because it is one the simplest and one of the fastest. Pluim et al. [3], showed that interpolation is the prevalent cause for the scalloping artifacts. They showed that it happens even in partial volume interpolation method introduced by Maes et al [4], and they concluded that sub-voxel accuracy as a result is unreasonable. Further analysis by Ji et al. [5] showed that sampling introduced also artifacts. More recently Rohde et al. [6;7] showed analytically that the scalloping artifacts are mostly due to uneven filtering by the interpolation process.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 43 – 49, 2006.

Several techniques have been proposed to reduce these artifacts. For example, Tsao [8] showed that they can be reduced by jittering the transformation parameters for each voxels before nearest neighbor interpolation, and by histogram blurring for the mutual information computation. Rohde et al. [6;7] found that sinc interpolation reduced scalloping. Over-sampling of the image and reduced binning of the joint histogram can also be beneficial [5].

We confirm in this communication that these scalloping artifacts are due to interpolation, and we propose a simple linear filter that completely removes them, smoothing the cross-correlation and mutual-information surfaces. In the next section we describe the method. In the following section we show experimental results on MR images. In the last section, we discuss the impact of this development on image registration.

## 2 Filter Design

Consider a discrete signal $s_j$ defined on a regularly spaced lattice $j \in \{1,...,N\}$. We would like to obtain the value of $s_i$ at a different location also defined over a regularly spaced lattice. Any arbitrary location can be written as a shift over an integer number of lattice grid plus a real number for in-between the lattice grid. We thus consider a shift by $\alpha \subset [0,1[$, with $i=j+\alpha$, and $j \in \square$ . After linear interpolation, the new signal $x_i$ can be written:

$$x_i = (1-\alpha)s_j + \alpha s_{j+1} \tag{1}$$

This equation can be interpreted as an adaptative linear filter: for $\alpha=0$ or 1, a simple shift of the lattice is needed, whereas when $\alpha=0.5$ for example, it corresponds to an averaging filter over two points. More complicated filters such as cubic or spline interpolation suffer from the same problem albeit to a lower extent [7].

Since linear interpolation corresponds to a filter that depends on the interpolation location, we suggest designing another filter that will also depend on the interpolation location, but with the opposite behavior. Such a filter, performed after linear interpolation, would result to a signal with constant variance. We use this criterion and design a constant variance filter for linear interpolation (cv-lin). To minimize the computation burden we chose the simplest non-causal filter:

$$y_i = ax_{i-1} + bx_i + ax_{i+1} \tag{2}$$

where $x$'s are samples on the new interpolated grid and $y$'s are values after application of cv-lin. We impose the normalization constraint

$$2a + b = 1 \text{ or } a = \frac{1-b}{2} \tag{3}$$

We specify $a$ and $b$ so that $\sigma_y^2 = \lambda \sigma_s^2$ with $\sigma_y^2$ and $\sigma_s^2$ the variance of $y$ and $s$ respectively and $\lambda \in \square_+$, a user-defined constant. We derived an equation for the variance following interpolation and solve to obtain a second-order equation for $b$ with solution below (derivation will be published elsewhere):

$$\hat{b} = \frac{(C-2D)+\sqrt{4D^2-2C^2+6\lambda C-8\lambda D}}{3C-4D}$$
$$\text{with: } \begin{cases} C = \left(2(\alpha-0.5)^2+0.5\right) \\ D = \alpha(1-\alpha) \end{cases} \tag{4}$$

The denominator exists for every $\alpha \subset [0,1]$, and the square root is positive for $\lambda > 1/3$. The value $\lambda = 1/3$ corresponds to the minimal variance achievable with a three-point averaging filter; i.e. [1 1 1]/3. Given the value of *b*, one can compute *a*, and specify the filter coefficients in (2). With $\lambda = 0.5$, the signal is left unchanged when it is the most filtered by linear interpolation (i.e. in the middle of grid points). By choosing $\lambda > 0.5$, cv-lin becomes a high-pass filter where the variance has been attenuated by linear interpolation more than 0.5; this might find application when image sharpness is desired. Extension to multidimensional data is straightforward because each dimension can be filtered independently. There will be as many $\alpha$'s as dimensions.

Note that for translation only $\alpha$ is the same for all pixels (one per dimension), but when rotation is present and for non-rigid body transformation in general, $\alpha$ will be different for each pixel, and therefore our proposed filter (Eq. 4) needs to be computed for each pixel (different coefficients a and b). The use of a lookup table to implement Eq. 4 improves significantly the computational cost.

## 3 Method

In order to test our filter we used multiple MR images from different anatomical parts. We show here typical results only on brain images because of space restriction. Since we would like to analyze the effects of interpolation, we cannot misregister an image because it would require an interpolation step. Instead we used the original image as the reference image and we interpolated the same image with different transformation parameters that included translation (horizontal and vertical) and rotation around the optimal parameters (all zero in this case). Cross-correlation and mutual-information were computed using standard equations [1;9]. Parameters were varied with a resolution of 0.03 pixels for translation and 0.03 degrees for rotation to assess sub-pixel perturbation. Similarity surfaces were plotted in three dimensions versus two parameters while one parameter was kept constant to its optimal value (zero). Similarity measures were also plotted versus each parameter individually while keeping the other two constant to their optimal values (zero).

Two separate cases were considered. First, raw images were used without any preprocessing. Second, images were filtered prior to the experiments with anisotropic diffusion filtering [10], and two separate noise realizations were added to each image to generate two different images. Noise standard deviation is expressed as a percentage of the dynamic range of each image.

## 4 Results

In Fig. 1, we show similarity curves for translation and rotation, when the floating image is an exact copy of the reference. Both cross-correlation and mutual-information show the characteristic scalloping artifact, which is removed with application of cv-lin. In this case, the global maximum always occurs at the correct location.

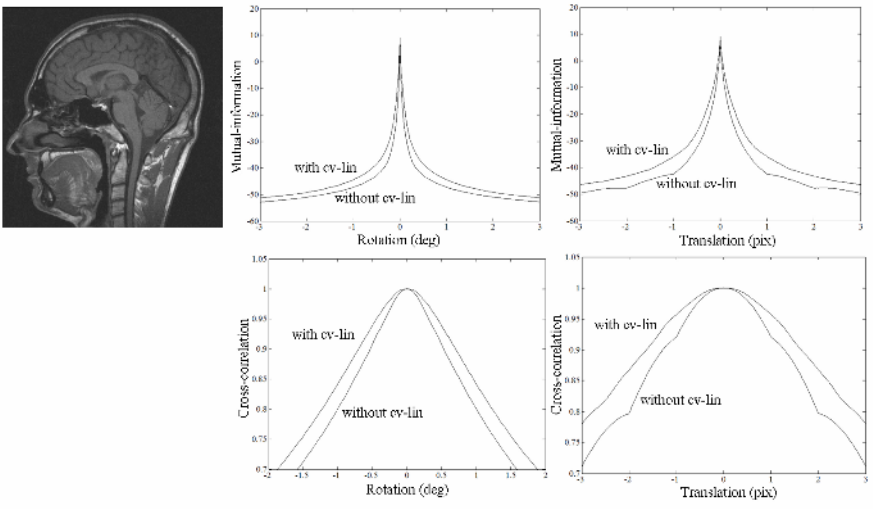



**Fig. 1.** Mutual information and cross-correlation as a function of translation and rotation. In this experiment the same image was used and misregistered to itself.

In Fig. 2 and 3, we used images having different noise realizations. In this case, noise realizations were added to a version of the original image filtered to reduce noise. Results are significantly different from those obtained in the previous experiment. Prior to cv-lin filtering, the global maximum does not occur at the true locations, and the scalloping artifacts are larger. Following application of cv-lin, the surface is smoothed and the erroneous biases of the global maxima are removed.

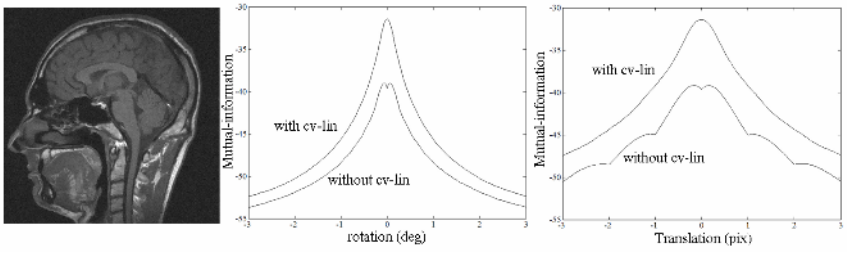



**Fig. 2.** We redid the same experiment as in Fig 1, but the images were first filtered and different noise realizations were added to yield two different images. Mutual information is shown for this 2% noise case. Note the global bias and local maxima have been canceled with cv-lin filtering.

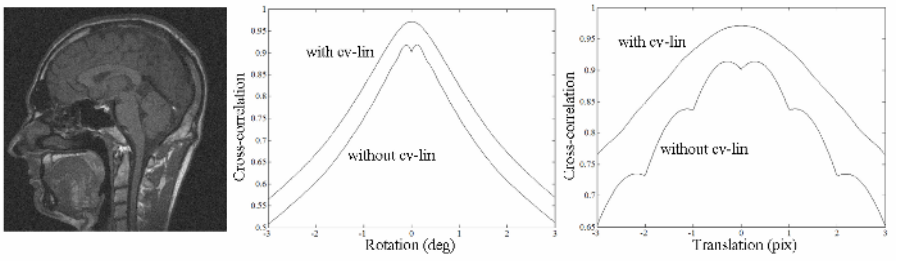



**Fig. 3.** As in Fig 2, the original image was first filtered and 4% noise was added. Cross-correlation is shown here as a function of rotation and translation.

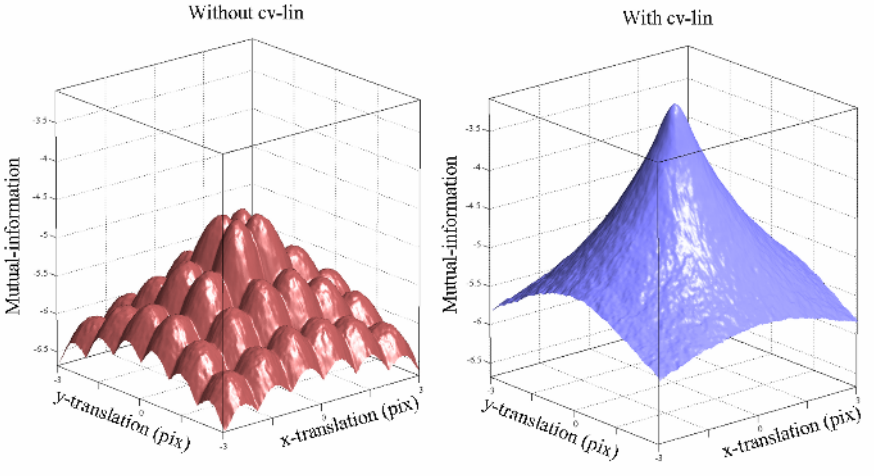



**Fig. 4.** Cross-correlation surfaces as a function of translation along x- and y-direction. The left panel shows the original surface with local maxima and global bias. In the right panel after cv-lin filtering the similarity surface is much smoother without noticeable interpolation-induced artifacts. Noise was 5%.

In Fig. 4 we show the cross-correlation surface as a function of translation in both directions to show how the proposed method efficiently removes the interpolation-induced artifacts.

## 5 Discussion

When a digital data set needs to be sampled at any arbitrary location, interpolation act as a spatially variable filter: when a new grid point matches an original one, the data is just shifted but otherwise unchanged; when a new grid point falls in between two original ones, the data is averaged. The interpolated data is thus unevenly filtered. As a consequence data features can be variably smoothed and the variance of the noise becomes variable across the data set. The effects on similarity measured commonly used in registration methods such as cross-correlation and mutual information are scalloping

artifacts creating local maxima and global bias. To address those issues we computed a simple filter that balances the effect of linear interpolation such that the data are evenly filtered across the data set.

The global bias of registration is reduced. If the optimal transformation happens to shift the new grid mostly in between the grid points of the original data, the interpolated data would be the most filtered. This results in sharpening of the histogram and the joint histogram increasing the mutual information measure. Compared to a slightly different transformation that would move the new grid closer to the original grid, almost no filtering would happen and the histograms would be unchanged. In this case the increase of MI due to the filtering could be higher than the increase of MI because the two data sets are better registered, and the global maximum of the MI would not happen for the optimal parameters. Similar explanation can be put forth for cross-correlation. Because cv-lin filters the data homogeneously such problem is much reduced as shown in the results section.

Local maxima are removed. The previous argument can be made again to explain the existence of local maxima in the similarity measures. When the transformation shifts the new grid in between the original grid, an increase in cross-correlation and the mutual information occurs. These local maxima can obviously trap the most sophisticated optimization algorithm. Other methods have been proposed to alleviate this problem as reviewed in the introduction. Some of them rely on filtering of all the data to minimize the artifacts with the risk to filter important data features that could be relevant for the registration. We think that cv-lin better addresses this problem because it filters the data only where it is needed. The partial-volume method introduced by Maes et al. [4] suffers also from interpolation artifacts as noted in [5] and [3], further investigation is needed to explore the benefit of cv-lin for this method.

Smoother similarity surfaces favorably impact registration methods. Sub-voxel accuracy is much improved in our preliminary tests (not shown) thanks to the removal of global bias. Convergence of optimization techniques is also improved because local maxima are reduced, thereby reducing cases where the optimization gets trapped, and also because a smoother surface increases the accuracy and robustness of derivatives and Hessian estimation. Since we used a simple linear filter, computation burden is limited and the new cv-lin filter can be used within existing registration algorithm without much increase in computation time. We are currently working on optimizing the implementation.

In conclusion, we have proposed a method to remove scalloping artifacts in cross-correlation and mutual-information. We designed a filter that produces constant variance in data that have been interpolated with linear interpolation. As a results global bias and local maxima are removed, which should improved registration methods based on the optimization of these similarity measures.

## References

1. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, Mutual-information-based registration of medical images: a survey, Medical Imaging, IEEE Transactions on, vol. 22, no. 8, pp. 986-1004, 2003.
2. M. Holden, D. L. Hill, E. R. Denton, J. M. Jarosz, T. C. Cox, T. Rohlfing, J. Goodey, and D. J. Hawkes, Voxel similarity measures for 3-D serial MR brain image registration, IEEE Trans. Med. Imaging, vol. 19, no. 2, pp. 94-102, Feb.2000.

3. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, Interpolation artefacts in mutual information-based image registration, Computer Vision and Image Understanding, vol. 77, no. 2, pp. 211-232, 2000.
4. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, Multimodality image registration by maximization of mutual information, IEEE Transactions on Medical Imaging, vol. 16, no. 2, pp. 187-198, Apr.1997.
5. J. X. Q. Ji, H. Pan, and Z. P. Liang, Further analysis of interpolation effects in mutual information-based image registration, IEEE Transactions on Medical Imaging, vol. 22, no. 9, pp. 1131-1140, 2003.
6. G. K. Rohde, A. S. Barnett, P. J. Basser, and C. Pierpaoli, Estimating intensity variance due to noise in registered images: Applications to diffusion tensor MRI, Neuroimage, vol. 26, no. 3, pp. 673-684, 2005.
7. G. K. Rohde, Berenstein C.A, and Healy D.M., Measuring image similarity in the presence of noise, in Medical Imaging 2005: Image Processing, 5747 ed. J. M. R. J.Michael Fitzpatrick, Ed. SPIE, 2005, pp. 132-143.
8. J. Tsao, Interpolation artifacts in multimodality image registration based on maximization of mutual information, IEEE Transactions on Medical Imaging, vol. 22, no. 7, pp. 854-864, 2003.
9. J. West, M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, M. L. Kessler, R. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. A. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire, M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods, Comparison and evaluation of retrospective intermodality brain image registration techniques, Journal of Computer Assisted Tomography, vol. 4, no. 21, pp. 554-566, 1997.
10. M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger, Robust anisotropic diffusion, Image Processing, IEEE Transactions on, vol. 7, no. 3, pp. 421-432, 1998.

# Symmetric Diffeomorphic Image Registration: Evaluating Automated Labeling of Elderly and Neurodegenerative Cortex and Frontal Lobe

Brian B. Avants, Murray Grossman, and James C. Gee

Penn Image Computing and Science Laboratory (PICSL)
University of Pennsylvania
Philadelphia, PA 19104-6389
avants@grasp.cis.upenn.edu

**Abstract.** One of the most challenging problems in modern neuroimaging is detailed characterization of neurodegeneration. Quantifying spatial and longitudinal atrophy patterns is an important component of this process. These spatiotemporal signals will aid in discriminating between related diseases, such as frontotemporal dementia (FTD) and Alzheimer's disease (AD), which manifest themselves in the same at-risk population. We evaluate a novel symmetric diffeomorphic image registration method for automatically providing detailed anatomical measurement over the aged and neurodegenerative brain. Our evaluation will compare gold standard, human segmentation with our method's atlas-based segmentation of the cerebral cortex, cerebellum and the frontal lobe. The new method compares favorably to an open-source, previously evaluated implementation of Thirion's Demons algorithm.

## 1 Introduction

Frontotemporal dementia (FTD) prevalence may be higher than previously thought and may rival Alzheimer's disease (AD) in individuals younger than 65 years [1]. Because FTD can be challenging to detect clinically, it is important to identify an objective method to support a clinical diagnosis. MRI studies of individual patients are difficult to interpret because of the wide range of acceptable, age-related atrophy in an older cohort susceptible to dementia. This has prompted MRI studies that look at both the rate and the anatomic distribution of change [2, 3].

Manual, expert delineation of image structures enables *in vivo* quantification of focal disease effects and serves as the basis for important studies of neurodegeneration [3]. Expert structural measurements from images also provide the gold-standard of anatomical evaluation. The manual approach remains, however, severely limited by the complexity of labeling $256^3$ or more voxels. Such labor is both time consuming and expensive to support, while the number of individual experts available for such tasks is limited. A third significant difficulty is the problem of inter-rater variability which limits the reliability of manual labeling [4]. While rarely available for large-scale data processing, an expert eye remains valuable for limited labeling tasks that give a basis for algorithmic evaluation.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 50–57, 2006.

Deformable image registration algorithms are capable of functioning effectively in time-sensitive clinical applications [5] and high-throughput environments and are used successfully for automated labeling and measurement research tasks. One challenge is reliable performance on non-standard data, as in studies of potentially severe neurodegenerative disorders. These types of images violate the basic assumptions of small deformations and simple intensity relationships used in many existing image registration methods.

Diffeomorphic image registration algorithms hold the promise of being able to deal successfully with both small and large deformation problems. State of the art methods also give full space-time optimizations, are symmetric with respect to image inputs and allow probabilistic similarity measures [6]. We will evaluate the performance of our symmetric diffeomorphic algorithm for high dimensional normalization of elderly and neurodegenerative cortical anatomy. We compare the method to the Demons algorithm which was shown to outperform other methods in a careful evaluation of inter-subject brain registration [7].

## 2 Methods

**Demons.** Thirion's Demons algorithm [8] is known to perform well in inter-subject deformable image registration. The method uses an elastic regularizer to solve an optical flow problem, where the "moving" image's level sets are brought into correspondence with those of a reference or "fixed" template image. In practice, the algorithm computes an optical flow term which is added to the total displacement (initially zero). The total displacement is then smoothed with a Gaussian filter. The process repeats for a set number of iterations for each resolution in a multi-resolution optimization scheme. The method is freely available in the Insight ToolKit and has been optimized by the ITK community (www.itk.org).

Dawant et al. used the Demons algorithm for segmenting the caudate nucleus, the brain and the cerebellum for a morphometric comparison of normal and chronic alcoholic individuals [9]. Their evaluation of the algorithm found reasonable agreement between automated and manual labeling. They also showed results on the automated labeling of hippocampus but did not evaluate performance. Their comparison used the kappa statistic (overlap ratio),

$$S(R1, R2) = \frac{2\sharp(R1 \cap R2)}{\sharp(R1) + \sharp(R2)}, \tag{1}$$

which measures both difference in size and location between two segmentations, $R1$ and $R2$. The $\sharp(R)$ operator counts the number of pixels in the region, $R$. This sensitive measure varies in the range $[0, 1]$ where values greater than 0.8 for smaller structures and 0.9 for larger structures are considered good.

**Symmetric Diffeomorphisms.** A diffeomorphism is a smooth, one-to-one, onto, invertible map. Shortest paths between elements in this space are termed *geodesic*. Diffeomorphic methods were introduced into medical computer vision

[10] for the purpose of providing a group theoretical, large deformation space-time image registration framework. Current developments in large deformation computational anatomy by Miller, Trouve and Younes extended the methods to include photometric variation and to use Euler-Lagrange equations [11]. However, these methods do not formulate the transformation symmetrically. They are only symmetric *in theory* and their implementation requires parallel computation [12].

Our current work extends the Lagrangian diffeomorphic registration technique described in [13]. This new formulation has symmetry properties required for a geodesic connecting two images, $I$ and $J$, in the space of diffeomorphic transformations. This formulation accounts for the natural symmetry in the problem: both images move along the shape (diffeomorphism) manifold. Symmetric diffeomorphisms guarantee two properties that are intrinsic to the notion of a geodesic path: the path from $I$ to $J$ is the same as it is when computed from $J$ to $I$, regardless of similarity metric or optimization parameters. Symmetry is required for distance estimates and makes results independent of arbitrary decisions about which image is "fixed" or "moving".

Our method is also unique in that it guarantees sub-pixel accurate, invertible transformations in the discrete domain. Driving forces may derive from landmark similarity with mutual information or other probabilistic measures of appearance relationships. This flexibility was inherited from our prior work [13]. Finally, the method is efficient enough to use on single-processor machines and in processing large datasets.

We define a diffeomorphism $\phi$ of domain $\Omega$, generally, for transforming image $I$ into a new coordinate system by $\phi I = I \circ \phi(\mathbf{x}, t = 1)$. The parameters of these transformations are time, $t$, a spatial coordinate, $\mathbf{x}$, and a velocity field, $\boldsymbol{v}(\mathbf{x})$ on $\Omega$, which is a square-integrable, continuous vector field [14]. The correspondence maps, $\phi$, are gained by integrating the velocity fields in time, $\phi(\mathbf{x}, 1) = \int_0^1 \boldsymbol{v}(\phi(\mathbf{x}, t))dt$; the distance is then $D(\phi(\mathbf{x}, 0), \phi(\mathbf{x}, 1)) = \int_0^1 \|\boldsymbol{v}\|_L dt$, where $L$ defines the linear operator regularizing the velocity. The functional norm, $\|\cdot\|_L$, induces regularity on the velocity field via linear differential operator $L = a\nabla^2 + b\mathbf{Id}$ ($a, b$ constants).

A basic fact of diffeomorphisms allows them to be decomposed into two parts, $\phi_1$ and $\phi_2$. We exploit this fact to define a variational energy that explicitly divides the image registration diffeomorphisms into two halves such that $I$ and $J$ contribute equally to the path and deformation is divided between them. This prior knowledge can be captured by including the constraint $D(\mathbf{Id}, \phi_1(\mathbf{x}, 0.5)) = D(\mathbf{Id}, \phi_2(\mathbf{z}, 0.5))$ directly in the optimization algorithm. The result is a method that finds correspondences with equal consideration of both images. Note that below we will derive the equations assuming intensity difference as a similarity measure, for simplicity. However, in actuality, we have a variety of statistical image similarity measures (robust intensity difference, cross-correlation, mutual information) at our disposal, as in [15], or employ user landmarks as in [13].

Define the image registration optimization time, $t \in [0, 1]$ where $t$ indexes both $\phi_1$ and $\phi_2$, though in opposite directions. The similarity seeks $\phi_1$ such that

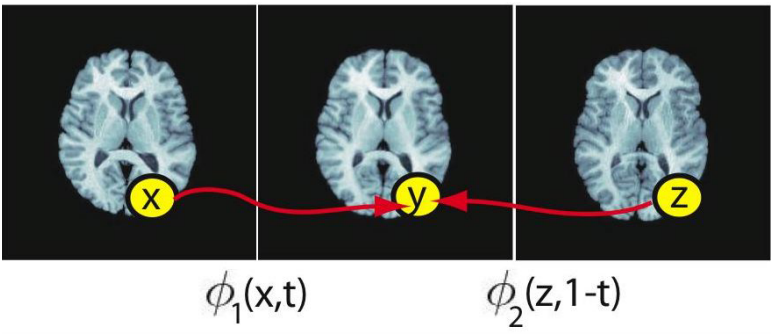



**Fig. 1.** An illustration of one curve in the SyN geodesic path between images. The image at far left and far right are the original images, $I$ and $J$. The maps, $\phi_1$ and $\phi_2$ are of equivalent length and map $I$ and $J$ to the mean shape between the images. The full path, $\phi$ and $\phi^{-1}$, are found by joining the paths $\phi_1$ and $\phi_2$.

$\phi_1(\mathbf{x}, 1)I = J$. Recalling the basic definition of diffeomorphisms allows us to write any geodesic through composing two parts. Then,

$$\begin{aligned} \phi_1(\mathbf{x}, 1)I &= J, \\ \phi_2^{-1}(\phi_1(\mathbf{x}, t), 1-t)I &= J, \\ \phi_2(\phi_2^{-1}(\phi_1(\mathbf{x}, t), 1-t), 1-t)I &= \phi_2(\mathbf{z}, 1-t)J, \\ \phi_1(\mathbf{x}, t)I &= \phi_2(\mathbf{z}, 1-t)J, \end{aligned} \tag{2}$$

converts the similarity term from $|\phi_1(\mathbf{x}, 1)I - J|$ to $|\phi_1(\mathbf{x}, t)I - \phi_2(\mathbf{z}, 1-t)J|^2$. A visualization of these components of $\phi$ is in Figure 1. The forward and backward optimization problem is then, solving to time $t = 0.5$,

$$\begin{aligned} E_{sym}(I, J) = \inf_{\phi_1} \inf_{\phi_2} \int_{t=0}^{0.5} \{ \|\boldsymbol{v}_1\|_L^2 + \|\boldsymbol{v}_2\|_L^2 + \\ \int_\Omega |I(\phi_1(t)) - J(\phi_2(1-t))|^2 d\Omega \} dt. \end{aligned}$$

$$\begin{aligned} &\text{Subject to:} \\ \boldsymbol{v}_1(0.5) = \boldsymbol{v}_2(0.5), &\|\boldsymbol{v}_1(t)\|_L^2 = \|\boldsymbol{v}_2(1-t)\|_L^2 \\ &\text{with each } \phi_i \in \mathit{Diff}_0 \text{ the solution of:} \\ &d\phi_i/dt = \boldsymbol{v}_i(\phi_i(t)) \text{ with } \phi_i(0) = \mathbf{Id}. \end{aligned} \tag{3}$$

Minimization with respect to $\phi_1$ and $\phi_2$, upholding the arc length constraint, provides the *symmetric* normalization (**SyN**) solution and also solves a 2-mean problem. Landmarks may also be included, as in the Lagrangian Push Forward method [13], by dividing the similarity term, as done with the image match terms above. This method is quite distinct from inverse consistent image registration (ICIR) [16] in which a variational term is used to estimate consistency. Furthermore, SyN provides an inverse that is guaranteed to be everywhere sub-pixel

accurate. This cannot be enforced by the ICIR formulation. Further details on the numerical methods employed in optimizing this energy may be found in [13].

**Implementation.** The Demons algorithm is freely available in the standard ITK distribution and has been quantitatively evaluated by the ITK community. We have implemented SyN within our extended version of the ITK deformable image registration framework, described in [17]. Therefore, we are in a position to measure performance gains by varying *only the transformation model*, as we use an identical similarity metric (optical flow). That is, for this study, SyN will use the itkDemonsRegistrationFunction as implemented in ITK for image forces. The only difference between the two methods that we compare is in the transformation model: we are therefore investigating if using our symmetric diffeomorphisms will enable better automated structure segmentations than the elastic model used by Demons.

**Dataset.** We now study the volumetric differences between elderly and frototemporal dementia cortex, with particular focus on the frontal lobe. We will compare the ability of two methods, SyN and Demons, to reproduce results gained from an expert user's labeling of our 20 image dataset. The frontal lobe is a major focus in research on aging, memory loss and dementia.

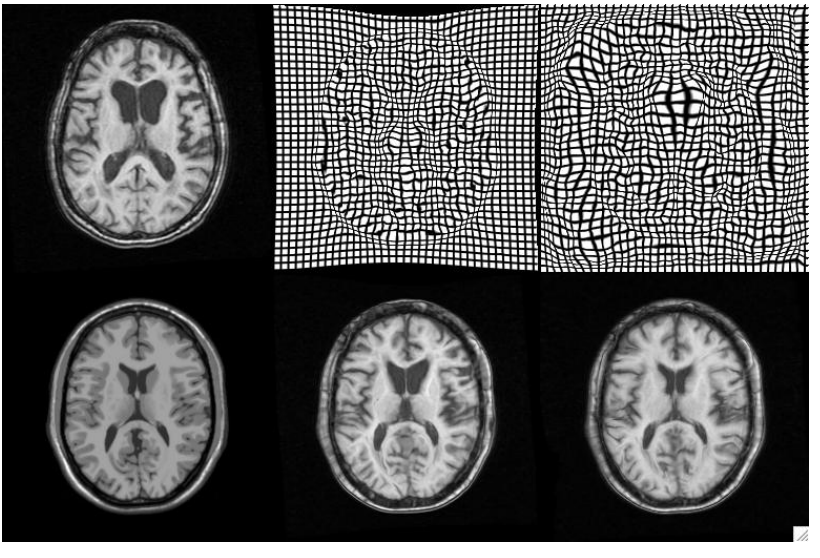


**Fig. 2.** The original FTD image, in upper left, was initially aligned to the atlas, lower left, via a rigid plus uniform scaling transformation. The subsequent Demons registration to the atlas, used for labeling, is in lower center. The corresponding grid deformation is in upper center. The SyN result is in lower right, while the corresponding grid deformation is upper right. The Demons method does a reasonable normalization, but leaves the ventricles and other smaller structures only partly normalized. The quadratic elastic penalty prevents the remaining shape differences from being captured.

We use a database of 20 T1 MRI images (0.85 x 0.85 x 1 mm, GE Horizon Echospeed 1.5 T scanner) from 10 normal elderly and 10 frontotemporal dementia patients. Each of the 20 images, along with the BrainWeb atlas, was manually labeled with the protocol described in [4]. This protocol was shown to be highly reproducible for both small and large structures via six-month intra-rater reliability and inter-rater reliability measurements. Left hippocampus labeling, for example, showed a 0.92 intra-rater overlap ratio (equation 1) and 0.83 average for inter-rater overlap. As the hippocampus is relatively small, these values are reasonable.

We compare the performance of the SyN algorithm to the Demons algorithm for automatically labeling this dataset. Both Demons and SyN were used to automatically segment the whole brain, cerebellum and frontal lobes by registering the labeled whole head MRI atlas to each individual whole head MRI. The atlas labelings are then warped by the same transformation into the space of the patient image. We then compute overlap ratios between the manual and automatic structural segmentations for each structure. An example comparison of the two methods is in figure 2.

## 3 Results and Discussion

Both algorithms produced segmentation results above the minimum threshold of 0.8 for all structures. A comparison between the average image produced after normalization of all images by each method is in figure 3. SyN had an average overlap ratio of 0.932 for cerebrum whereas the Demons value was 0.919; for frontal lobe, SyN mean = 0.901 while Demons = 0.882; for cerebellum, SyN mean = 0.883 while Demons = 0.861. We computed Student's T-test to evaluate whether SyN outperforms Demons for labeling these structures. SyN produces statistically significant better ($T > 2.5$ ) results over the whole dataset for each of these structures: frontal lobe ($P < 0.03$), cerebellum ($P < 0.04$) and cerebrum ($P < 0.016$). The gap in performance on frontal lobe ($P < 0.018$) and cerebrum ($P < 0.003$) increases when we focus only on the FTD results. This separation is


**Fig. 3.** A section of the template image is at left. The average of all elderly images registered to the template by Syn and Demons are shown next, followed by the associated error images (intensity difference between the average and the template). The Demons average appears to have slightly larger ventricles and fewer sulci, as a result of a general trend toward underestimating the deformation. This effect is exaggerated, for Demons, when FTD data is included.

caused by the presence of larger deformation in the FTD subjects some of which may not be captured by Demons. Note that, in the past, similar evaluations have shown Demons to outperform other methods [7].

**Clinical Significance.** We will now compare the ability of the Demons and SyN results to estimate the difference in structural volume given by the manual outlining. We achieve this by using the outlined frontal lobe on the atlas image to mask each jacobian, $\mathcal{J}_i$, as computed from each registration. The volume of the frontal lobe for individual $i$ is estimated from registration by $\sum$ (all $\mathbf{x}$ in the atlas frontal lobe) $\mathcal{J}_i(\mathbf{x})$. We therefore have three distinct measurements of the elderly and frontotemporal dementia frontal lobe: one from the expert labeling, one from the Demons algorithm and one automated measure from our SyN method. Next, we quantify the ability of the regularized, automated normalization methods to approximate the results gained from the labor intensive manual approach.

FTD and aging both negatively impact memory. However, memory loss in individuals with FTD is severe and accelerated. We expect this clinical presentation to be reflected in cortical structure, particularly, with FTD, in the frontal lobe. Here, we focus on lobar volume, as opposed to shape.

If we test for significant differences in size as given by manual labeling, we find that elderly frontal lobes are larger with a significance of $P < 0.0003$, a very strong result. The Demons method *does NOT show significant results*, with $P < 0.088$. The SyN method, on the other hand, *does show significant results* with $P < 0.030$, where all significance values are assessed with permutation testing. Thus, we can see how an apparently small difference in performance (as measured by overlap ratio) can have an impact on the study outcome. Note that the Demons and SyN significance are both smaller than the manual results. This is caused by segmentation bias towards the normal atlas (which has a more regular labeling) and the fact that registration-based segmentations are smoothed (Demons more than SyN), while the manual segmentations are not. Note that Demons smooths the global transformation but does not restrict to the space of diffeomorphisms. SyN explicitly smooths only the velocity field, but restricts to diffeomorphisms, a space which can capture large deformation differences as occurring in elderly and FTD data.

**Conclusion.** This preliminary comparison shows the distinct advantage of SyN for segmenting elderly and neurodegenerative cerebrum, cerebellum and frontal lobe. Note that, in addition to better performance, SyN provides a dense space-time map and transformation inverses. The differences in performance are consistent, statistically significant and have a major impact on study outcome. One can extrapolate even larger differences between SyN and algorithms with lower dimensionality than either Demons or SyN. For this reason, along with the theoretical advantages that translate into practical benefits, we promote diffeomorphic algorithms in neuroimaging research, in particular when studying non-standard datasets, such as FTD and AD.

## References


1. Ratnavalli, E., Brayne, C., Dawson, K., Hodges, J.: The prevalence of frontotemporal dementia. Neurology **58** (2002) 1585–1586
2. Fox, N., Crum, W., Scahill, R., Stevens, J., Janssen, J., Rossor, M.: Imaging of onset and progression of alzheimer's disease with voxel-compression mapping of serial magnetic resonance images. Lancet **358** (2001) 201–205
3. Studholme, C., Cardenas, V., Blumenfeld, R., Schuff, N., Rosen, H.J., Miller, B., Weiner, M.: Deformation tensor morphometry of semantic dementia with quantitative validation. Neuroimage **21** (2004) 1387–1398
4. Sparks, B., Friedman, S., Shaw, D., Aylward, E., Echelard, D., Artru, A., Maravilla, K., Giedd, J., Munson, J., Dawson, G., Dager, S.: Brain structural abnormalities in young children with autism spectrum disorder. Neurology **59** (2002) 184–92
5. Dawant, B., Li, R., Cetinkaya, E., Kao, C., , Fitzpatrick, J., Konrad, P.: Computerized atlas-guided positioning of deep brain simulators: A feasibility study. In Gee, J., Maintz, J.B., eds.: Workshop on Biomedical Image Registration, Philadelphia (July, 2003) 142–150
6. Avants, B., Epstein, C.L., Gee, J.C.: Geodesic image interpolation: Parameterizing and interpolating spatiotemporal images. In: ICCV Workshop on Variational and Level Set Methods. (2005) 247–258 in press.
7. Hellier, P., Barillot, C., Corouge, I., Gibaud, B., Le Goualher, G., Collins, D., Evans, A., Malandain, G., Ayache, N., Christensen, G., Johnson, H.: Retrospective evaluation of inter-subject brain registration. IEEE Transactions on Medical Imaging **22** (2003) 1120–1130
8. Thirion, J.: Non-rigid matching using demons. In: IEEE Computer Vision and Pattern Recognition. (1996) 245–251
9. Dawant, B., Hartmann, S., Thirion, J.P., Maes, F., Vandermeulen, D., Demaerel, P.: Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations, part II: methodology and validation on severely atrophied brains. IEEE Trans Med Imaging **18** (1999) 971–926
10. Trouve, A.: Diffeomorphism groups and pattern matching in image analysis. Intl. J. Comp. Vis. **28** (1998) 213–221
11. Miller, M., Trouve, A., Younes, L.: On the metrics and Euler-Lagrange equations of computational anatomy. Annu. Rev. Biomed. Eng. **4** (2002) 375–405
12. Beg, F., Miller, M., Trouve, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int. J. Comp. Vision **61** (2005) 139–157
13. Avants, B., Schoenemann, P.T., Gee, J.C.: Landmark and intensity-driven lagrangian frame diffeomorphic image registration: Application to structurally and functionally based inter-species comparison. Medical Image Analysis (2005) e-pub, in press.
14. Arnold, V.I.: Ordinary Differential Equations. Springer-Verlag: Berlin (1991)
15. Hermosillo, G., Chefd'Hotel, C., Faugeras, O.: A variational approach to multimodal image matching. Intl. J. Comp. Vis. **50** (2002) 329–343
16. Johnson, H.J., Christensen, G.E.: Consistent landmark and intensity-based image registration. IEEE Trans. Med. Imaging **21** (2002) 450–461
17. Yoo, T.: Insight Into Images: Principles and Practice for Segmentation, Registration and Image Analysis. AK Peters, Ltd., Natick, MA (2003)

# Deformation Based Morphometry Analysis of Serial Magnetic Resonance Images of Mouse Brains

Satheesh Maheswaran[1], Hervé Barjat[4], Simon Bate[4], Thomas Hartkens[3], Derek L.G. Hill[3], Michael F. James[4], Lorna Tilling[4], Neil Upton[4], Jo Hajnal[2], and Daniel Rueckert[1]

[1] Dept. of Computing, 180 Queen's Gate, South Kensington Campus, Imperial College, London, SW7 2AZ
satheesh@doc.ic.ac.uk
[2] Imaging Sciences Department, Imperial College, Hammersmith Hospital, London, W12 0NN
[3] IXICO Ltd, London, NW1 0NH
[4] Neurology and GI Center of Excellence for Drug Discovery, GlaxoSmithKline, Harlow, Essex, CM19 5AW

**Abstract.** Deformation based morphometry is used to detect differences in in-vivo Magnetic Resonance Image (MRI) of the mouse brain obtained from two transgenic strains: TASTPM mice that over-express proteins associated with Alzheimer's disease, and wild-type mice. MRI was carried out at four time points. We compare two different methods to detect group differences in the longitudinal and cross-sectional data. Both methods are based on non-rigid registration of the images to a mouse brain atlas. The whole brain volume measurements on 27 TASTPM and wild-type animals are reproducible to within 0.4% of whole brain volume. The agreement between different methods for measuring volumes in a serial study is shown. The ability to quantify changes in growth between strains in whole brain, hippocampus and cerebral cortex is demonstrated.

## 1 Introduction

Mouse models are used in many biomedical research areas to study issues ranging from development to drug efficacy. The models are made popular by their rapid life cycle in comparison to other animals as well as the wealth of genetic information and the technology available to modify them. The human [1] and mouse genome [2] project has made a large amount of data available on genes, which makes our ability to understand disease processes a real possibility. MRI allows the 3D morphology of anatomical structures to be examined in-vivo. Recent efforts on the use of MRI to study anatomical differences between mouse strains include [7], [8]. Chen et al [6] reported on significant differences of neuroanatomy in three different mouse strains. Kovacevic et al [4] reported on a variational MRI atlas constructed from nine excised mice brains. Verma et al [5]

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 58–65, 2006.

reported that there were sharp contrast between tissue anisotropy changes in the cortex and in major white-matter fibers during the first 80 post-natal days.

To the best of our knowledge, all existing registration methods have been applied to ex-vivo mouse brain images. However, the use of ex-vivo images does not allow serial imaging studies which are aimed at detecting differences across time within an individual (e.g. atrophy) or within a group. In this paper we focus on developing methods for MRI of the brain acquired in-vivo. This raises new challenges for both the computational techniques as well as the image acquisition techniques used in in-vivo studies. In this paper, we demonstrate that existing image registration techniques are capable of producing acceptable segmentations on in-vivo mouse brains. The longitudinal volumes can be measured in two different ways, one way is to register everything to a common reference and the other is to use each animal's first scan as its control. We explore the two methods and report on their level of consistency. Finally, we show the ability of our method to quantify changes in growth between strains in anatomical structures.

## 2 Materials and Methods

We have developed image registration techniques which have been applied to breast images, cardiac motion tracking and brain development in neonates [3]. The global registration between two images is modelled by an affine transformation and the local registration is described by a free-form deformation (FFD) based on B-splines. The FFD is modelled as a mesh of control points. Normalised mutual information is used on a voxel-based similarity measure which is insensitive to intensity changes. The registration is achieved by minimising the cost function, which represents a combination of the cost associated with the smoothness of the transformation and the cost associated with the image similarity [3].

### 2.1 Data

Animal experiments complied with GSK ethical and UK legal requirements. Transgenic (Tg) mice were used which over-expressed proteins linked to Alzheimer's disease [9]. The transgene leads to abundant deposition of the protein $\beta - amyloid$ in the brain from approximately 5 months of age. Wild-type (Wt) mice, which behaves as the $C57Bl6$ background strain, were used as controls. MRI was carried out at $6, 9, 11$ and 14 months of age.

### 2.2 MR Acquisition Protocol

Prior to MR imaging the mice were anesthetized with isoflurane and their heads were immobilised in a custom-built head holder. Their core temperature and spontaneous respiration was monitored during the imaging process and kept constant. After recovery from anesthesia the mice were returned to their home cages. Images were acquired on a 4.7T Brucker Biospec $40cm$ horizontal bore magnet

using 25$mm$ diameter volume receive and 12 cm diameter volume transmit coils. A multi-slice (120 slices) multi-echo CPMG dataset was obtained using interleaved scans of slice thickness 0.31mm, with final voxel resolution $78\times78\times156\mu m$, was acquired at each time-point. The field of view was $20 \times 20 \times 18.6mm$ and the matrix was $256 \times 256$. Total imaging time was 2 hours.

During data acquisition, we acquired two interleaved volumes off set by a half a slice thickness (the sequences $S_1$ and $S_2$). This acquisition was repeated to obtain another image of the same animal (the sequences $S_3$ and $S_4$). The sequences were interleaved to achieve better pseudo resolutions. An image of an animal was obtained by merging $S_1$ and $S_2$. Two images may be obtained of the same animal by merging $S_1, S_2$ and $S_3, S_4$. During these acquisitions the animals were left in the scanner and not moved. The $S_3$ and $S_4$ sequences were acquired in the event animal movement during $S_1$ and $S_2$. This provides data to check for reproducibility in measurements as well the ability to deal with potential loss of data due to animal motion during scanning.

### 2.3 Measuring Volumes In-Vivo in Longitudinal Data

Our aim is to carry out volumetric analysis of the brain and its structures over time. In order to achieve our goal, we require a mouse atlas. We use the LONI atlas [10] from the Mouse Atlas Project (MAP) as our atlas for segmentation. All 27 structures labelled in the LONI atlas were propagated into each animals native space. Since the MR acquisition parameters for the LONI atlas and our MR acquisitions were different, we register the LONI atlas to one of our animals, to obtain a segmentation of the brain. This segmentation is then used as the atlas in our study. Figure 1 shows the mouse brain atlas annotated with 5 structures for illustration purposes.

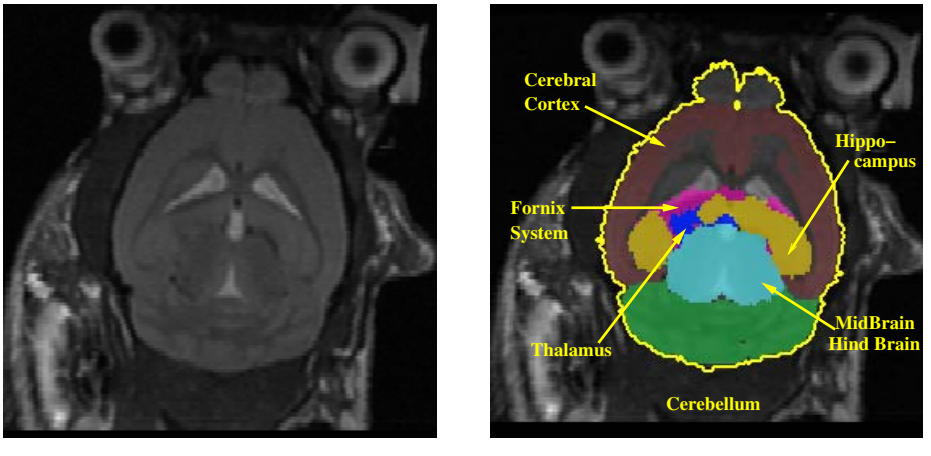



**Fig. 1.** The annotated atlas obtained via registering to the LONI atlas

**Segmentation by Direct Atlas Registration.** In the first stage the brain atlas is aligned to each image using an affine registration. Using the affine registration, an approximate segmentation of the brain for each animal is obtained. To compensate for residual misalignment after affine registration, the segmentation is then dilated by 5 voxels. The purpose of the approximate segmentation of the brain was to improve the speed and accuracy of the subsequent non-rigid registration. In the second stage, we perform a non-rigid registration of the mouse brain atlas to each image using the approximate brain mask obtained previously. We use a multi-resolution approach starting with $2mm$ control point spacing going down to $0.25mm$. After non-rigid registration, the labels of the brain atlas, $\mathcal{O}_{atlas}$, are then propagated into the space of each animal to obtain a final segmentation. We can then compute the growth rate as the difference in volume across time within the same animal as shown below:

$$\Delta V_t(\mathcal{O}_{label}) = V_{t+1}(\mathcal{O}_{label}) - V_t(\mathcal{O}_{label})$$

**Segmentation by Indirect Atlas Registration.** In the previous method, each image is segmented by registering the mouse brain atlas directly to each image. An alternative method for segmentation is to register the images of an animal to a baseline scan of this animal. In our example, the first time point at 6 months serves as a baseline and the images acquired at 9, 11 and 14 months are registered to the baseline image using non-rigid registration. The baseline image can be segmented using the method described in the previous section. The growth rate for each structure, $\mathcal{O}_{label}$, can be computed by integrating the determinant of the Jacobian matrix, J, of the transformation between time points across every voxel, x, belonging to $\mathcal{O}_{label}$ in the baseline image as shown below:

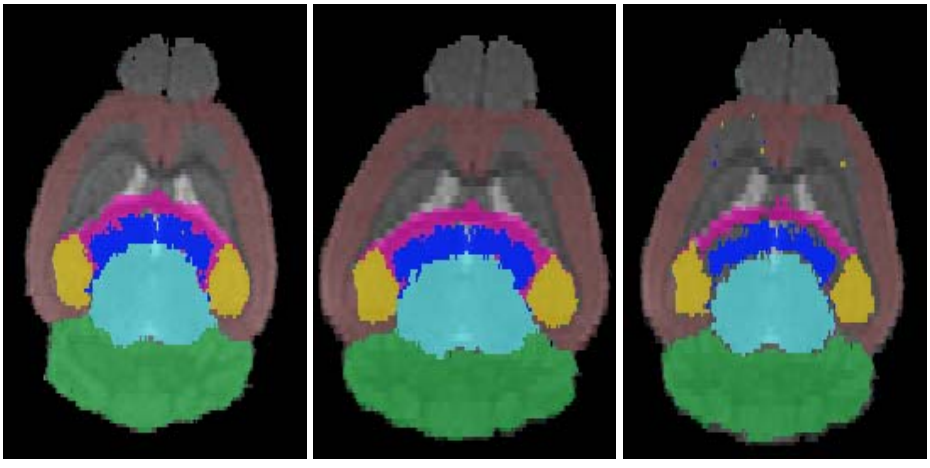
$$\Delta V_t(\mathcal{O}_{label}) = \int_{x \in \mathcal{O}_{baseline}} det|J(x)|$$

## 3 Results

In this section, we present the results using the two registration methods described above. All 27 labels were propagated from the atlas using direct and indirect atlas registration methoods. Figure 2 shows an illustration of the segmentations obtained using direct and indirect atlas registration methods for an animal at 9 months. The propagated labels include, MidBrain-HindBrain(MB-HB), hippocampus, thalamus, fornix system, cerebral cortex and cerebellum.

### 3.1 Consistency Measure

We wanted to investigate the consistency of our label propagation methods. In order to calculate the consistency in our segmentations, we transform all segmentations from both registration methods in to a common reference (atlas space). The Similarity Index(SI) is then calculated between the atlas and segmentaion



(a) Atlas (b) Direct registration (c) Indirect Atlas Registration

**Fig. 2.** Label propagation using direct and indirect atlas registration. Fig 2(b) shows segmentation obtained via direct registration and Fig 2(c) shows segmentation obtained via indirect atlas registration.

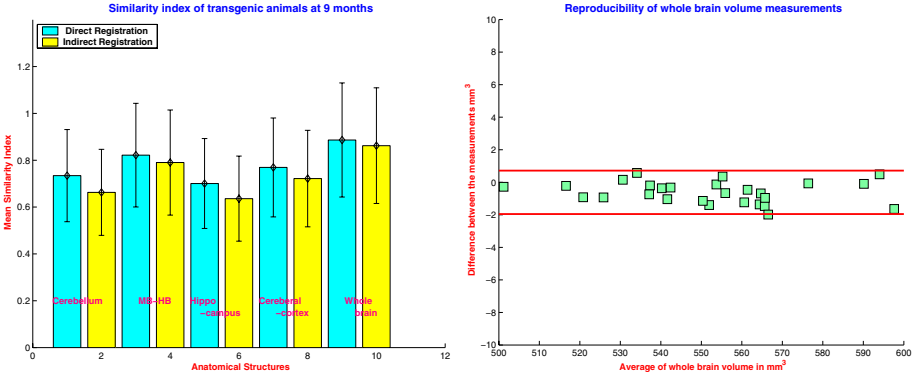
obtained via label propagation. The SIs were calculated for all 9 month old transgenic animals. The metric used to calculate the SI is given below

$$SI = \frac{|I_{atlas} \cup I_{segmentation}|}{I_{atlas} \cap I_{segmentation}}$$

Figure 3(a) shows the SI for both methods. We measured the SI for whole brain, hippocampus, cerebellum, MB-HB and cerebral cortex. The error bars shows the standard deviations of the SI measurements within groups. As can be seen from Figure 3(a), we failed to observe any significant differences in the SI between the two methods described above. However, since indirect registration requires more registrations to obtain a segmentation relative to the direct registration method, it is more likely to have the larger registration errors of the two. Therefore we use the direct registration method as the preferred label propagation method for all of our experiments from here on.

### 3.2 Reproducibility of Volume Measurements

In order to investigate the ability of our registrations to reproduce the measurements, we perform the same registration on both images obtained via merging $S_1S_2$ and $S_3S_4$ sequences acquired at 6 months using the direct registration method. Figure 3(b) shows a Bland-Altman plot of the measurements for the two sets of data for whole brain volume. As can be seen from Figure 3(b) we are able to reproduce our results to within 0.4% of whole brain volume.

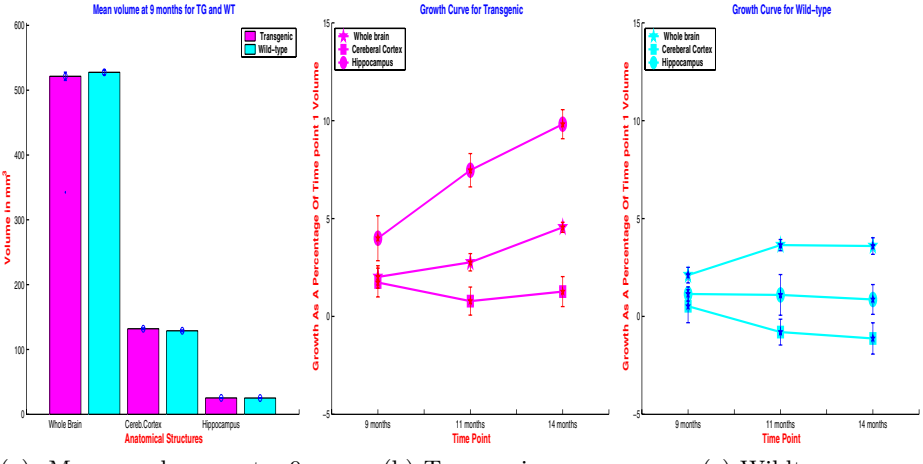



(a) Consistency Measure (b) Reproducibility

**Fig. 3.** Fig 3(a) shows consistency measure of whole brain and anatomical structures using direct and indirect registration. Fig 3(b) shows Bland-Altman plot of repeated measurements for whole brain volumes using direct registration method.

### 3.3 Measuring Growth Rates

We use the direct registration method to measure growth rates for the whole brain, as well as anatomical structures like hippocampus and cerebral cortex as defined by the LONI atlas. The growth curves for transgenic and wild-type are shown in Figure 4(b) and 4(c) respectively. The graph on the y-axis




(a) Mean volume at 9 Months (b) Transgenic (c) Wildtype

**Fig. 4.** Growth measurements for Transgenic and Wildtype Mouse brains. Growth is plotted as a percentage of volume change with respect to time point 1 measurement. Fig4(a) shows the mean volume for TASTPM and WT.

represents growth as a percentage of the measured volume at 6 months. The x-axis represents time as discrete points. The error bars shows the group standard deviations of the measurements at a given time point. The growth curves shown in Figure 4(b) and 4(c) provides evidence of growth in structures for both strains of animals. More over our results shows different cross-sectional growth rates for whole brain, cerebral cortex and hippocampus. However there exists a larger group variation of structural volumes within the transgenic in comparison to the wild-type. This variation of measurement is likely to be due to the biological variability of the genetically modified groups to the controls. Figure4(a) shows the mean volume at 9 months for transgenic and wild-type.

## 4 Discussion and Conclusions

We set out to measure longitudinal volume changes in mouse brains using image registration techniques. To the best of our knowledge, all existing work has been done on ex-vivo brains. Our data set consisted of two strains of animals (transgenic and wild-type) imaged at $6, 9, 11$ and 14 months. In order to analyse our registration method's abilities of reproducibility, we measured volumes repeatedly on the same animals using two different images of the same animal. Direct and indirect registration methods were used to measure volumes. We found no significant differnce in measured volumes between the direct and indirect registration methods. We were able to quantify growth in between strains which were statistically significant.

In conclusion, we have shown that our registration methods allows both cross-sectional and longitudinal studies. We are able to reproduce our whole brain volume measurements to within 0.4%. The SI values shows a high similarity between the atlas and the segmentations obtained via registration. We have also demonstrated the ability to quantify changes in growth between strains in whole brain, hippocampus and cerebral cortex. In future, we plan to focus on analysing the diffusion and T2 maps as well as applying sophisticated data mining techniques.

## References

1. International Human Genome Sequencing Consortium: Initial sequencing and analysis of the human genome. Nature, **409**, 860–921, 2001
2. Mouse Genome Sequencing Consortium: Initial sequencing and comparative analysis of the mouse genome. Nature, **420**, 520–562, 2002
3. D.Rueckert and L. I. Sonoda and C. Hayes and D. L. G. Hill and M. O. Leach and D. J. Hawkes: Nonrigid Registration Using Free-Form Deformations : Applications to Breast MR Images. IEEE Transactions On Medical Imaaging, **18**, 8, 712–721, 1999
4. N. Kovacevic and J. T. Henderson and E. Chan and N. Lifshitz and J. Bishop and A.C. Evans and R.M. Henkelmen and X. J. Chen: A Three-dimensional MRI Atlas of the Mouse Brain with Estimates of the Average and Variability. Cerebral Cortex, **15**, 5, 639–645, 2005

5. R. Verma and S. Mori and D. Shen and P. Yarowsky andJ. Zhang and C.Davatzikos: Spatiotemporal maturation patterns of murine brain quantified by diffusion tensor MRI and deformation-based morphometry. PNAS, **102**, 19, 6978–6983, 2005
6. X.J. Chen and N. Kovacevic and N. Lobaugh and J.G. Sled and R.M.Henkelman and J.T.Henderson: Neuroanatomical differences between mouse strains as shown by high-resolution 3D MRI. NeuroImage, **29**, 99–105, 2005
7. A.A. Ali and A.M. Dale and A.B. Badea and G.A.Johnson: Automated segmentation of neuroanatomical structures in multispectral MR microscopy of the mouse brain. NeuroImage, **27**, 425–435, 2005
8. B.J. Nieman and N.A.Bock and J.Bishop and X.J. Chen and J.G. Sled and J. Rossant and R.M. Henkelman: Magnetic resonance imaging for detection and analysis of mouse phenotype. NMR Biomed., **18**, 447–468, 2005
9. D.R. Howlett and J.C.Richardson and A.Austin and AA Parsons, and S.T. Bate and D.C. Davies and M.I. Gonzalez: Cognetive correlates of $A\beta$ deposition in male and female mice bearing amyloid precursor protein and presenilin-1 mutant trasngenes. Brain Research., **1017**, 130–136, 2004
10. A. MacKenzie-Graham, E. Lee, I. D. Dinov, M. Bota, D. W. Shattuck, S. Ruffins, H. Yuan, F. Konstantinidis, A. Pitiot, Y. Ding, G. Hu, R. E. Jacobs and A.W. Toga: A multimodal, multidimensional atlas of the C57BL/ 6J mouse brain. J. Anat., **204**, 93–102, 2004

# Canonical Correlation Analysis of Sub-cortical Brain Structures Using Non-rigid Registration

Anil Rao[1], Kola Babalola[2], and Daniel Rueckert[1]

[1] Visual Information Processing Group, Department of Computing,
Imperial College London, 180 Queen's Gate, London SW7 2BZ, U.K
[2] Division of Image Science & Bio-medical Engineering,
University of Manchester, Stopford Building, Oxford Road, Manchester, M13 9PTS, U.K

**Abstract.** In this paper, we present the application of canonical correlation analysis to investigate how the shapes of different structures within the brain vary statistically relative to each other. Canonical correlation analysis is a multivariate statistical technique which extracts and quantifies correlated behaviour between two sets of vector variables. Firstly, we perform non-rigid image registration of 93 sets of 3D MR images to build sets of surfaces and correspondences for sub-cortical structures in the brain. Canonical correlation analysis is then used to extract and quantify correlated behaviour in the shapes of each pair of surfaces. The results show that correlations are strongest between neighbouring structures and reveal symmetry in the correlation strengths for the left and right sides of the brain.

## 1 Introduction

The area of computational anatomy is a rapidly developing discipline [14]. With the increasing resolution of anatomical scans of the human brain, a number of computational approaches for characterising differences in the shape and neuro-anatomical configuration of different brains have emerged. Morphometric techniques can be classified into techniques that deal with differences in brain shape (deformation-based morphometry [3, 9]) and those which deal with differences in the local composition of brain tissue after removing global shape differences (voxel-based morphometry [1]). Even though both approaches require warping of images into a standard reference space using either elastic [20, 4, 13] or fluid [8, 7, 6] registration techniques, they differ fundamentally in the way the resulting deformation fields are used. In deformation-based morphometry the deformation fields themselves are used to study similarities and differences, while in voxel-based morphometry these fields are used principally for normalisation. There is currently an active discussion in the neuroscience community regarding the advantages and disadvantages of both methods [5, 2].

A prominent example of modelling the variability of neuro-anatomical structures across a population is the probabilistic atlas of the human brain developed at the Montreal Neurological Institute (MNI) [10] where MR images from 305 subjects were mapped into stereotactic space, intensity normalised and averaged on a voxel-by-voxel basis as part of the International Consortium for Brain Mapping (ICBM) [19]. An alternative approach uses statistical models such as active shape models [12] or active

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 66–74, 2006.

appearance models [11] to represent the variability of the anatomy across subjects. Statistical shape modelling refers to the analysis of the shapes of sub-structures (such as the lateral ventricles in the centre of the brain) and aims to describe their variation across subjects and between groups of subjects (e.g., comparing ventricle size and shape between Alzheimer's sufferers and age-matched normals). Work on shape modelling is limited by the generally unsolved problems of how to segment these structures and determine correspondences across subjects, and also by the relatively unexplored area of how different structures vary statistically relative to each other within the brain. The latter problem will be addressed specifically in this paper.

In this paper we describe the application of canonical correlation analysis for the analysis of the inter-structure shape variation within the brain. Canonical correlation analysis (CCA) is a multivariate statistical tool for describing and quantifying correlated variation between sets of vector variables. It is an extension of multilinear regression and has been used to analyse data in a number of different application areas. Within the field of imaging, canonical correlation analysis has been previously used in image segmentation of magnetic resonance spectroscopic images [16] and the identification of noise in functional magnetic resonance images [21]. Canonical correlation analysis has also been used to estimate the shapes of obscured anatomical sections of the brain from visible structures in magnetic resonance images [17]. However, there it was used as a predictive tool for a limited number of structures within the brain. Here we use it to extract highly correlated factors (or modes) of variation in shape between a number of different anatomical structures within the brain and an associated correlation coefficient that quantifies the degree of correlation in this shape variation. This reveals statistical dependencies between different shapes in the brain that ultimately we would like to incorporate into a hierarchical model-fitting scheme. In the next section we describe the mathematical formulation of canonical correlation analysis before presenting results of the application of this technique in the analysis and prediction of brain structures in section 3.

## 2 Canonical Correlation Analysis

The object of canonical correlation analysis is to extract and quantify correlations between two sets of vector variables, $\mathbf{X} = \{\boldsymbol{x}_i\}, \mathbf{Y} = \{\boldsymbol{y}_i\}$. The technique determines linear combinations of the components of the vector variables in $\mathbf{X}$ that are maximally correlated with linear combinations of the components in $\mathbf{Y}$, and the strength of each of the correlations is described by a corresponding correlation coefficient that lies between zero and one. The linear combinations, known as the canonical modes, give insight into the relationships between the two sets of variables [18].

The canonical modes $\hat{\boldsymbol{a}}_k, \hat{\boldsymbol{b}}_k$ and correlation coefficients $\rho_k$ for $\mathbf{X}$ and $\mathbf{Y}$ are calculated by solving the eigenvalue equations

$$\begin{aligned} C_{XX}^{-1} C_{XY} C_{YY}^{-1} C_{XY}^{T} \hat{\boldsymbol{a}}_k &= {\rho_k}^2 \hat{\boldsymbol{a}}_k \\ C_{YY}^{-1} C_{XY}^{T} C_{XX}^{-1} C_{XY} \hat{\boldsymbol{b}}_k &= {\rho_k}^2 \hat{\boldsymbol{b}}_k \end{aligned} \tag{1}$$

where $C_{XX}$, $C_{YY}$ and $C_{XY}$ are the covariance matrices describing variation within $\mathbf{X}$, $\mathbf{Y}$, and between $\mathbf{X}$ and $\mathbf{Y}$ respectively [16]. The calculated modes $\hat{\boldsymbol{a}}_k$ and $\hat{\boldsymbol{b}}_k$ are then the

linear combinations of variable components in $\mathbf{X}$ and $\mathbf{Y}$ respectively that have a corresponding correlation coefficient $\rho_k$. The number of modes and correlation coefficients determined by a canonical correlation analysis of $\mathbf{X}$ and $\mathbf{Y}$ will be equal to the minimum of the number of dimensions in the vectors $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$. A single correlation coefficient $\rho$ representing the overall correlation between $\mathbf{X}$ and $\mathbf{Y}$ can be determined by averaging the correlation coefficients over all calculated canonical modes.

Canonical correlation analysis has certain maximal properties similar to those of principal components analysis (PCA). However, the two techniques differ fundamentally in that while CCA focuses on relationships *between* two groups of variables, PCA considers interrelationships *within* a single group of variables [18]. If, for example, we were to pool the two sets of variables $\mathbf{X} = \{\boldsymbol{x}_i\}$ and $\mathbf{Y} = \{\boldsymbol{y}_i\}$ into a single set and then perform a PCA, we would lose the distinction between the two sets of data as the PCA does not 'know' to which data set each variable originated from. The resulting modes would then just model the variation of the composite data set without explicitly describing the dependencies of the individual data sets on each another.

## 3 Canonical Correlation Analysis of Brain MR Data

### 3.1 Method

A set of MR images of 93 subjects from the Centre for Morphometric Analysis (CMA), Boston, was used to create a training set of surfaces over which the canonical correlation analysis was applied. The images were obtained at resolution 1mm x 1.5mm x 1mm and had been manually labelled in order to delineate structures within the brain by experts at the CMA. Firstly, a reference subject was chosen and the surfaces of 17 different sub-cortical brain structures of this reference subject were calculated from its labelled image. These surfaces are shown in figure 1 and represent the left and right lateral ventricle, left and right caudate, left and right putamen, left and right accumbens, left and right pallidum, left and right thalamus, left and right amygdala, left and right hippocampus and the brain stem.

In order to model the variation in the surfaces of these structures across all subjects, correspondences between each reference surface point and the corresponding surface in each of the other subjects must be determined. These were calculated by registering the labelled magnetic resonance images of each subject to the reference image using a B-Spline FFD registration algorithm which represents each transformation as a sum of a global affine component and a non-rigid component based on B-Splines [20]. The optimal transformation is found by maximising the label consistency of the labelled images which measures the degree of their alignment.

The registrations were then used to create a set of surface points for each of the structures over all 93 subjects. Firstly, for a given non-reference subject, the corresponding reference to subject transformation was applied to the reference surfaces to give a set of subject surface points. Then, for each structure, the calculated surface points of the 93 subjects were Procustes-aligned using scaling, rotations and translations [15, 12] to give the final surface point coordinates for each structure and subject. The Procustes-alignment ensures that any subsequent statistical modelling of the generated surface

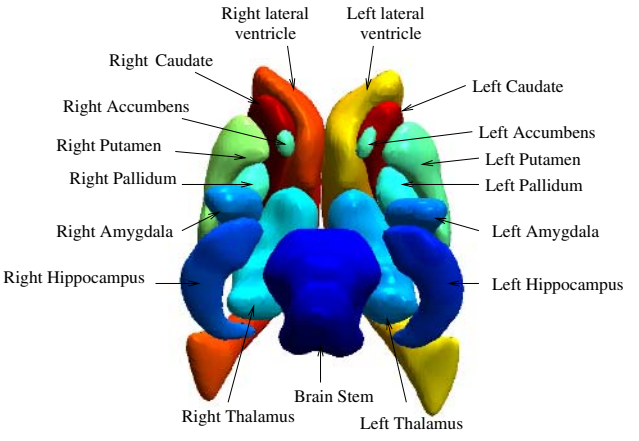



**Fig. 1.** This colour figure shows each of the 17 structures that we considered

points only describes variations in the shape of a structure over the training data rather than variation due to differences in the position, orientation or size of a structure.

For each individual structure X, the vectors $\boldsymbol{x}_i$ representing the surface point coordinates for structure X of the $i$th subject $0 < i < 93$, were pooled to form a set of vectors $\mathbf{X} = \{\boldsymbol{x}_i\}$. Prior to performing the canonical correlation analysis of surface point coordinates, a principal components analysis was performed on the surface point coordinates for each individual structure across the training data to reduce the dimensionality of the data. The dimension reduction minimizes the computational memory burdens of the canonical correlation analysis and also eliminates colinearity in the data which can cause instability in the calculation of CCA. Fifty-five modes of variation were retained from the prinicipal components analysis of each structure ensuring that at least 95% of the variation in that structure across the training data could be represented. For each structure X, the set $\mathbf{X} = \{\boldsymbol{x}_i\}$ was then transformed into its corresponding principal components basis to give a new set of vectors $\tilde{\mathbf{X}} = \{\tilde{\boldsymbol{x}}_i\}$. Canonical correlation analysis was then performed for all pairs of structures X and Y using the corresponding sets of vectors $\tilde{\mathbf{X}} = \{\tilde{\boldsymbol{x}}_i\}$ and $\tilde{\mathbf{Y}} = \{\tilde{\boldsymbol{y}}_i\}$. In each case, fifty-five canonical modes and correlations are determined describing the correlated behaviour between structure X and structure Y. For a given pair of structures X and Y, these correlations were then averaged to give a final correlation coefficient $\rho$ between zero and one describing the strength of the correlations between the two structures. We chose not to retain the actual canonical modes as we were interested in quantifying correlations rather than analyzing the correlated behaviour itself.

### 3.2 Results

Figure 2 shows the canonical correlation coefficients for each pair of structures as a grey-level matrix image in which brighter areas represent higher correlations. We can see that the correlation coefficients achieve the maximum value of one along the top-left to bottom-right diagonal as this line represents the correlations between a structure

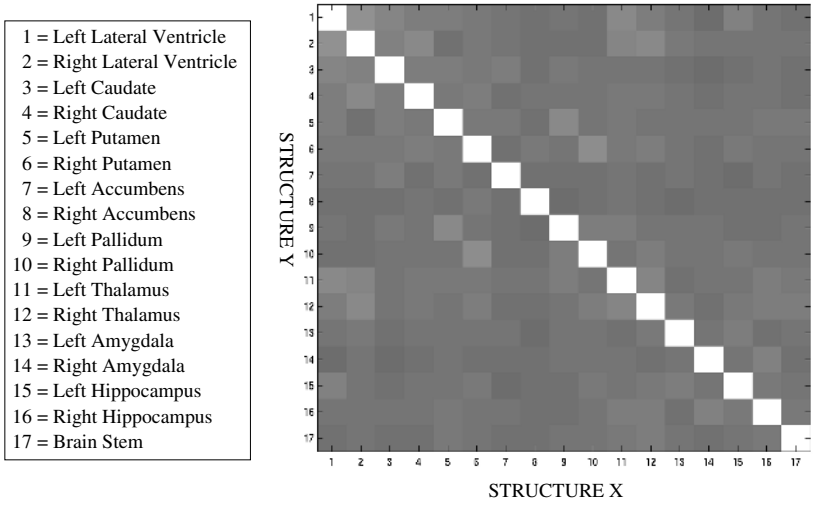


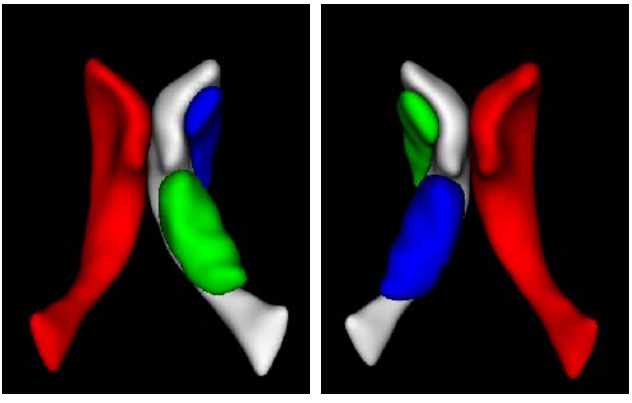

**Fig. 2.** This figure shows the matrix image of the correlations $\rho$ between each pair of structures. In the image, bright areas represent strong correlations close to 1, while dark areas represent weaker correlations close to 0.5.

**Table 1.** This table lists each structure and its most correlated structure

| **Structure** | **Most Correlated Structure** | $\rho$ | **Structure** | **Most Correlated Structure** | $\rho$ |
|---|---|---|---|---|---|
| L. Lateral Ventricle | R. Lateral Ventricle | 0.7836 | L. Pallidum | L. Putamen | 0.7714 |
| R. Lateral Ventricle | L. Lateral Ventricle | 0.7836 | R. Pallidum | R. Putamen | 0.7748 |
| L. Caudate | L. Lateral Ventricle | 0.7604 | L. Thalamus | L. Lateral Ventricle | 0.7659 |
| R. Caudate | R. Lateral Ventricle | 0.7703 | R. Thalamus | R. Lateral Ventricle | 0.7681 |
| L. Putamen | L. Pallidum | 0.7714 | L. Amygdala | L. Hippocampus | 0.7492 |
| R. Putamen | R. Pallidum | 0.7748 | R. Amygdala | R. Hippocampus | 0.7509 |
| L. Accumbens | L. Caudate | 0.7437 | L. Hippocampus | L. Lateral Ventricle | 0.7509 |
| R. Accumbens | R. Putamen | 0.7406 | R. Hippocampus | R. Amygdala | 0.7509 |
| Brain Stem | R. Thalamus | 0.7483 | | | |

and itself which are always perfect. The matrix image is also symmetrical about this diagonal as the calculation of the correlation coefficients between any two structures is independent of which is taken to be structure X and which is taken to be structure Y.

If we consider those correlations lying off the leading diagonal we can see that each pair of structures is correlated to different degrees. For example, the right pallidum is better correlated to the right putamen ($\rho = 0.7748$) than to the left accumbens ($\rho = 0.7236$). This means that, across the training data, the shape of the right pallidum varies in a more correlated fashion with the shape of the right putamen than with the shape of the left accumbens. In table 1 we show the best correlated structure (strongest correlate) for each of the 17 structures and we can see that the strongest correlates of



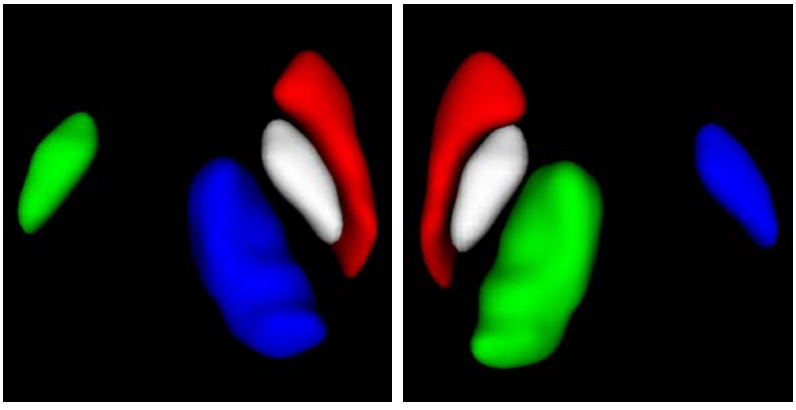
(a) (b)

**Fig. 3.** This figure shows each of the lateral ventricles in grey with the three most strongly correlated structures to each of them in in red, green and blue in order of decreasing correlation strength. Figure (a) shows the left lateral ventricle (grey), right lateral ventricle (red), left thalamus (green) and left caudate (blue) and figure (b) shows the right lateral ventricle (grey),the left lateral ventricle (red), the right caudate (green) and the right thalamus (blue).

each structure are proximal to it. For example, the left/right caudates are most strongly correlated with the left/right lateral ventricles respectively and we can see in figure 1 that these structures are next to each other. Similarly, each of the lateral ventricles, putamen, accumbens, pallidum, thalamus and amygdala are best correlated with proximal structures. Intuitively, this makes sense as one would expect variation in the shape of a structure to be reflected in the shapes of proximal and neighbouring structures. An example of this relationship between proximity and correlation strength is visualised in figure 3, in which the three strongest correlates for the left/right lateral ventricles are shown. In this figure the lateral ventricles are shown in grey and the correlated structures are shown in red, green and blue in order of decreasing correlation strength.

There is also a degree of symmetry in the strongest correlates for structures that appear in both sides of the brain. For example, the left putamen is most strongly correlated with the left pallidum, while conversely the *right* putamen is most strongly correlated with the *right* pallidum. This symmetry is repeated for all the structures apart from the hippocampus and accumbens. Such an example is shown in figure 4, in which the three most correlated structures to the left and right pallidum are shown. Here we can see that the strongest correlates of each pallidum are the neighbouring putamen and that the pallidum are strongly correlated to each other. We find that equivalent structures on each side of the brain are relatively highly correlated for all the structures apart from the putamen and accumbens.

In order to investigate correlations that are not associated with symmetries in the brain, we also performed a canonical correlation analysis in which equivalent structures on each side of the brain were concatenated into single structures. The resulting surfaces were Procustes aligned and 56 principal components of the PCA were retained before performing the canonical correlation analysis. Table 2 shows the strongest and



(a) (b)

**Fig. 4.** This figure shows the left and right pallidum in grey with the three most strongly correlated structures to each of them in red, green and blue in order of decreasing correlation strength. Figure (a) shows the left pallidum (grey), left putamen (red), right pallidum (green) and left thalamus (blue) and figure (b) shows the right pallidum (grey), right putamen (red), right thalamus (green) and left pallidum (blue).

**Table 2.** This table lists each composite structure and its most/least correlated composite structure

| **Structure** | **Most/Least Correlated Structure** | $\rho$ |
|---|---|---|
| Lateral Ventricle | Caudate/Accumbens | 0.7857/0.7367 |
| Caudate | Lateral Ventricle/Brain Stem | 0.7857/0.7352 |
| Putamen | Pallidum/Hippocampus | 0.7829/0.7387 |
| Accumbens | Caudate/Brain Stem | 0.7489/0.7346 |
| Pallidum | Putamen/Hippocampus | 0.7829/0.7343 |
| Thalamus | Lateral Ventricle/Amygdala | 0.7848/0.7427 |
| Amygdala | Hippocampus/Brain Stem | 0.7613/0.7248 |
| Hippocampus | Amygdala/Pallidum | 0.7613/0.7343 |
| Brain stem | Thalamus/Amygdala | 0.7563/0.7248 |

weakest correlates for each of the 9 composite structures. As would be expected, the strongest correlations depicted in table 2 match those shown in table 1 for the individual structures. Overall, the strongest correlates to each of the 9 structures are proximal to them, while the weakest correlates of structures are distal to them in all cases apart from the lateral ventricles.

## 4 Discussion

The results of the canonical correlation analysis imply that there are differing degrees of correlated variation between the shapes of different structures within the brain. The CCA gives larger correlations for structures that are close to each other, which

suggests that the shapes of structures both influence, and are influenced by, those structures that are proximal to them. Intuitively, this makes sense as one would expect variation in the shape of a structure to be reflected in the shapes of proximal and neighbouring structures. However, it should be noted that the correlations and associated modes determined by canonical correlation analysis do not necessarily describe a large amount of the *variation* between structures but instead describe the most *correlated* behaviour in that variation.

The most interesting aspect of this work that we are currently pursuing is the incorporation of canonical correlation analysis into a hierarchical model-fitting algorithm. Since CCA quantifies the strengths of the correlations between the shapes of different structures, it can be combined with shape prediction techniques such as partial least squares regression to guide a model fitting in a hierarchical fashion. Such a technique would involve performing CCA on individual structures and groups of structures as we have described in this paper.

## References


1. J. Ashburner and K. J. Friston, *Voxel-based morphometry – the methods*, NeuroImage **11** (2000), no. 6, 805–821.
2. ———, *Why voxel-based morphometry should be used*, NeuroImage **14** (2001), no. 6, 1238–1243.
3. J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price, and K. Friston, *Identifying global anatomical differences: Deformation-based morphometry*, Human Brain Mapping **6** (1998), 638–57.
4. R. Bajcsy and S. Kovačič, *Multiresolution elastic matching*, Computer Vision, Graphics and Image Processing **46** (1989), 1–21.
5. F. L. Bookstein, *"Voxel-based morphometry" should not be used with imperfectly registered images*, NeuroImage **14** (2001), no. 6, 1452–1462.
6. M. Bro-Nielsen and C. Gramkow, *Fast fluid registration of medical images*, Proc. 4th International Conference Visualization in Biomedical Computing (VBC'96), 1996, pp. 267–276.
7. G. E. Christensen, S. C. Joshi, and M. I. Miller, *Individualizing anatomical atlases of the head*, Proc. 4th International Conference Visualization in Biomedical Computing (VBC'96), 1996, pp. 434–348.
8. G. E. Christensen, M. I. Miller, J. L. Mars, and M. W. Vannier, *Automatic analysis of medical images using a deformable textbook*, Computer Assisted Radiology (Berlin, Germany), Springer-Verlag, June 1995, pp. 146–151.
9. M. K. Chung, K. J. Worsley, T. Paus, D. L. Collins C. Cherif, J. N. Giedd, J. L. Rapoport, and A. C. Evans, *A unified statistical approach to deformation-based morphometry*, NeuroImage **14** (2001), no. 3, 595–606.
10. D. L. Collins, P. Neelin, T. M. Peters, and A. C. Evans, *Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space*, Journal of Computer Assisted Tomography **18** (1994), no. 2, 192–205.
11. T. F. Cootes, G. J. Edwards, and C. J. Taylor, *Active appearance models*, Proc. 5th European Conference on Computer Vision (ECCV'98), vol. 2, 1998, pp. 484–498.
12. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, *Active Shape Models - their training and application*, Computer Vision and Image Understanding **61** (1995), no. 1, 38–59.
13. J. Gee, M. Reivich, and R. Bajcsy, *Elastically deforming 3D atlas to match anatomical brain images*, Journal of Computer Assisted Tomography **17** (1993), no. 2, 225–236.

14. U. Grenander and M. I. Miller, *Computational anatomy: An emerging discipline*, Quarterly of Applied Mathematics **56** (1998), no. 4, 617–694.
15. B. Horn, *Closed-form solution of absolute orientation using unit quaternions*, Journal of the Optical Society of America **4** (1987), 629–642.
16. T. Laudadio, P. Pels, L. Lathauwer, P. Hecke, and S. Huffel, *Tissue segmentation and classification of mrsi data using canonical correlation analysis*, Magnetic Resonance in Medicine **54** (2005), 1519–1529.
17. T. Liu, D. Shen, and C. Davatzikos, *Predictive modeling of anatomic structures using canonical correlation analysis*, IEEE International Symposium on Biomedical Imaging, 2004.
18. K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate analysis*, Academic Press, Belfast, U. K., 1982.
19. J. Mazziotta, A. Toga, A. Evans, P. Fox, and J. Lancaster, *A probabilistic atlas of the human brain: Theory and rationale for its developement. the international consortium for brain mapping*, NeuroImage **2** (1995), no. 2, 89–101.
20. 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** (1999), no. 8, 712–721.
21. L. Zollei, L. Panych, E. Grimson, and W. Wells, *Exploratory identification of cardiac noise in fmri images*, Sixth Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI '03), Lecture Notes in Computer Science, 2003.

# A Novel 3D/2D Correspondence Building Method for Anatomy-Based Registration

Guoyan Zheng

MEM Research Center, University of Bern, Stauffacherstrasse 78, CH-3014, Bern, Switzerland
Guoyan.Zheng@MEMcenter.unibe.ch

**Abstract.** The application of fluoroscopic images in operation is pervasive, especially for orthopaedic surgery. Anatomy-based 3D/2D registration, rigid or non-rigid, has been proven to improve the accuracy and precision of various image-guided therapies. One of the key steps for a successful anatomy-based registration is to establish 3D/2D correspondence between the 3D model and the 2D images. This paper presents a novel 3D/2D correspondence building method based on a non-rigid 2D point matching process, which iteratively uses a symmetric injective nearest-neighbor mapping operator and 2D thin-plate spline based deformation to find a fraction of best matched 2D point pairs between features detected from the X-ray images and those extracted from the 3D model. The estimated point pairs are further ranked by their shape context matching cost and those with high cost are eliminated. The remaining point pairs are then used to set up a set of 3D point pairs such that we turn a 3D/2D registration problem to a 3D/3D one, whose solutions are well studied. Rigid and non-rigid registration algorithms incorporating the novel 3D/2D correspondence building method are presented. Quantitative and qualitative evaluation results are given, which demonstrate the validity of our method.

## 1 Introduction

The application of fluoroscopic images in operation is pervasive, especially for orthopaedic surgery. Disadvantages of fluoroscopy include two-dimensional (2D) projection image from single view, limited field of view, distorted image, and high radiation to both the patient and the surgical team. Various papers [1, 2] have described methods of calibration and registration of fluoroscopic images using an optical localizer, thus allowing to compute the position of the surgical tools relative to the patient anatomy with respect to acquired images during intervention. However, the surgeon still needs to mentally fuse projection images taken from different view points. No real three-dimensional (3D) information is available.

One way to address this problem is to do anatomy-based 3D/2D registration, rigid or non-rigid, using 3D model extracted from either CT data or from statistical shape models. The co-registered or reconstructed anatomical model can then provide detailed 3D information for the considered bone structure, which has been proven to improve the accuracy and precision of various image-guided therapies [3, 4].

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 75 – 83, 2006.

However, anatomy-based 3D/2D registration is not a trivial task. One of the key steps for a successful 3D/2D registration is to establish correspondence between the 3D model and the 2D images. This paper presents a novel 3D/2D correspondence building method based on a non-rigid 2D point matching process, which iteratively uses a symmetric injective nearest-neighbor mapping operator and 2D thin-plate spline (TPS) based deformation to find a fraction of best matched 2D point pairs between features detected from the 2D images and those extracted from the 3D models. The estimated point pairs are further ranked by their shape context matching cost and those with high matching cost are eliminated. The remaining point pairs are then used to set up a set of 3D point pairs such that we turn a 3D/2D registration problem to a 3D/3D one, whose solutions are well studied.

### 1.1 Our Contributions

The main contribution of this paper is the novel 3D/2D correspondence building method based on the iterative non-rigid point matching process and on the shape context distance based outlier elimination. This matching process belongs to a more general non-rigid point matching framework [5, 6]. Thin-plate spline (TPS) is used in their papers and the present paper as the parameterization of the deformation. The main difference between theirs and ours are: (1) the design philosophy is different. The algorithms presented in [5, 6] are targeted for more general applications, whereas our method is specially designed for establishing 3D/2D correspondence. Our emphasis is to find a small number of point pairs with "good" matching quality; (2) their solutions to 2D point correspondence are different from ours.

Compared with the previously published 3D/2D correspondence building methods, our method differs in the way to extract the apparent contours and in the features extracted from the input images. Unlike in [3, 7, 8, 9] where the apparent contours are extracted in the primary space by an exhausting searching of all edges on the surface modes, we use a method introduced by Hertzmann and Zorin [10], which is based on dual space theory and provides a faster and smoother contour generator. More importantly, our 3D/2D correspondence building method is based directly on edge pixels, which can be easily extracted by applying edge detectors to the input images. In contrast, all previously published methods require an explicit contour extraction, which may be quite difficult when the shapes involved become complex or when the backgrounds of the images become complex.

## 2 Image Feature Extraction

A standard implementation of Canny edge detector with hysteresis [11] is used to find the edge pixels of the considered bone structure from the input images. To suppress spurious edges, Gaussian convolution kernel with large width is used to smooth the input images first. The detected edge pixels are further processed using the knowledge about the fluoroscope. Detected edge pixel whose intensity is below a pre-selected threshold or whose distance to the image center is bigger than a certain threshold is eliminated.

## 3 Apparent Contour Extraction

For fast extraction a smooth apparent contour from a given surface model, we use the approach introduced by Hertzmann and Zorin [10]. This approach first constructs a dual representation of the given surface mesh in a four-dimensional (4D) space based on the position and tangent planes of every vertex. The focal point's dual (a plane in 4D) intersects with the mesh triangles dual. Before hand, the approach normalizes the dual vertices using the $l_\infty$ norm so that the vertices end up on one of the unit hypercube's sides. This reduces the problem to intersecting the triangles on a hypercube's sides with the focal point's dual plane, i.e., to intersect triangles in eight 3D unit cubes (the eight hypercube sides) with a plane. By using Octree for each hypercube sides, the extraction process can be greatly speeded up.

For a point $\Omega_j^s, (j = 1,2,\cdots,M)$ on the extracted apparent contours using the projection parameters of input X-ray image *s*, we do a forward projection of this point onto image *s* to get its 2D position $A_j^s$. Each 2D point in $\{A_j^s\}$ is then associated to a 3D point in $\{\Omega_j^s\}$. In the following section, we will describe an iterative matching process for build 2D association between the points in $\{A_j^s\}$ and the detected edge pixels in the X-ray image *s*.

## 4 Iterative Non-rigid 2D Matching Process

Following the general framework of point matching in [5], we also formularize the 2D matching as a two-stage process: correspondence and spatial mapping. TPS is used here for parameterization of the spatial mapping. But unlike [5, 6], we solve the correspondence problem differently. To make the description simple, we denote the detected edge pixels in image s as $I^s = \{I_i^s, i = 1,2,\cdots,N\}$ and the 2D projection of the apparent contours as $A^s = \{A_j^s, j = 1,2,\cdots,M\}$. Here we focus on 2D matching in one image. The overall correspondence is established by combining matched point pairs found in all input images.

**Definition 1: Injective nearest-neighbor (IN).** A point $I_i^s$ can only be matched to at most one point of $A^s$ and this point must be the closest one in $A^s$ to $I_i^s$.

**Definition 2: Cross-matching.** Assume there are two matched pairs $(I_i^s, A_j^s)$ and $(I_m^s, A_n^s)$. If the line segment from $I_i^s$ to $A_j^s$ intersects at a point with the line segment from $I_m^s$ to $A_n^s$, we define this event as cross-matching.

**Definition 3: Symmetric Injective nearest-neighbor (SIN).** $I_i^s$ and $A_j^s$ is a symmetric injective nearest-neighbor if and only if $A_j^s$ is the closest point in $A^s$ to $I_i^s$ and $I_i^s$ is the closest point in $I^s$ to $A_j^s$.

**Definition 4: Symmetric injective nearest-neighbor mapping operator (SIN-MO).** We define the action of finding a number of SIN's from two points set as symmetric inject nearest-neighbor mapping operator.

**Claim 1:** If we apply SIN-MO on two point sets to find a number of matched point pairs, all of them are one-to-one mapping.
**Claim 2:** If we apply SIN-MO on two point sets to find a number of matched point pairs, there is no cross-matching (see Appendix for a proof).

The overall iterative non-rigid 2D matching process can be described as follows.

**Input:** Two point sets $I^s$ and $A^s$, and a weight parameter $\lambda$
**Output:** A list of 2D matched point pairs
**Initialization:** We first calculate the centroids of $I^s$ and $A^s$, and then translate $A^s$ so that its centroid is aligned with the centroid of $I^s$
**Iteration:** It is a dual update process taking those points $I^s$ as references.
**Stage 1: Update the correspondence:** Applying SIN-MO on $I^s$ and $A^s$ to find a number of SIN's. Let's denote the set of SIN's as $\{(I_a^s, A_a^s); a = 1,2,\cdots,K\}$

**Stage 2: Update the positions of all points in $A^s$:** This is done in a two-step procedure.

A. Compute a 2D TPS-based spatial mapping $f$ using the estimated set of SIN's by minimizing the following cost function:

$$E_{2D-TPS}(f) = \sum_{a}^{K} \| I_a^s - f(A_a^s) \|^2 + \lambda \iint [(\frac{\partial^2 f}{\partial x^2})^2 + (\frac{\partial^2 f}{\partial x \partial y})^2 + (\frac{\partial^2 f}{\partial y^2})^2] dxdy \quad (1)$$

B. Update the positions of all points in $A^s$ based on the estimated TPS transformation $f$

Repeat stage 1 and 2 a certain times (e.g. 30) or until convergence.

## 5 Using Shape Context Distance to Improve the Robustness

The 2D matching process described above has the advantages of robustness to certain outliers and of automatic exclusion of cross matching, which is an important property for preservation of topology in non-rigid registration. However, it is possible for the proposed algorithm to create false matches. In this paper, we use the shape context matching cost to further rank the point pairs estimated by the 2D matching process and to eliminate those with high cost such that the false matching rate is reduced.

The shape context of a point is a measure of the distribution of other points relative to it [12]. Consider two points, $p_i$ in one shape and $q_j$ in the other shape. Their shape contexts are $h_i(k)$ and $h_j(k)$, two $K$-bin normalized histograms at $p_i$ and $q_j$ for $k$ = 1, 2, …, $K$, respectively. Let $C_{ij}$ denote the cost of matching these two points. As shape context are distributions represented as histograms, the $\chi^2$ test statistic is used to define $C_{ij}$ in [12] as following equation:

$$C_{ij} = \frac{1}{2}\sum_{k=1}^{K}\frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \tag{2}$$

Those point pairs extracted by the 2D matching process are regarded as "best matched" in terms of shortest distance in the context of iteratively deforming one shape to match the other. By adding a step of shape context matching cost checking, we also take the neighborhood distributions of those point pairs into consideration.

## 6 3D Matched Point Pair Building

Assume that we have found a set of 2D matched point pairs $\{(I_b^s, A_b^s); b = 1, 2, \cdots, L\}$, we are trying to build the corresponding 3D point pairs in this step as follows. For a 2D point $I_b^s$, we can find a projection ray $r_b^s$ emitting from the focal point of image $s$ through point $I_b^s$. Additionally, for its matched point $A_b^s$, we always know its associated 3D point $\Omega_b^s$ on the apparent contour of the model whose projection onto the image $s$ is $A_b^s$. By computing a point $v_b^s$ on the ray $r_b^s$ which has the shortest distance to $\Omega_b^s$, we can build a 3D point pair $(v_b^s, \Omega_b^s)$. Combining all these 3D point pairs, we can establish 3D/2D correspondence between the input model and images.

## 7 Anatomy-Based Rigid or Non-rigid Registrations

As soon as a set of matched 3D point pairs are established, we have turned a 3D/2D registration problem to a 3D/3D one whose solutions are well studied. A complete description of these solutions is beyond the scope of this paper. Here we would like to present two algorithms to illustrate how to incorporate the proposed correspondence building method into rigid or non-rigid 3D/2D registrations.

### 7.1 Rigid 3D/2D Registration

The corresponding rigid 3D/3D registration problem is a well-known problem and several efforts have been made to solve it. One of the most popular methods is the *iterative closest point* (ICP) algorithm [13, 14]. The ICP is based on the search of pairs of closest points, and the computation of a paired-point matching transformation. The result transformation is then applied to one set of points, and the procedure is iterated until convergence. The local converging behavior of ICP algorithm requires a proper initialization, which can be achieved by a paired-point matching using anatomical landmarks of the considered bone structures.

Incorporating the present 3D/2D correspondence building method, we turn a 3D/2D registration problem to a 3D/3D one. Similar to the ICP algorithm, in each step we calculate a paired-point matching transformation based on the estimated point

pairs and then apply the transformation to one side of the registration. This procedure is then iterated until convergence. We call this procedure the *Iterative Best Matched Projection Point* (IBMPP) algorithm. The difference between the IBMPP algorithm and the ICP algorithm is that in each step point sets from both sides of the registration are changeable for the IBMPP algorithm while only one point set is changeable for the ICP algorithm.

### 7.2 Non-rigid 3D/2D Registration

Recently, statistical shape models based non-rigid 3D/2D registration has drawn a lot of attentions [7, 8, 9, 15]. Basically there are two ways to optimize the solutions: (1) do it directly on the images [15]; and (2) convert it to a non-rigid 3D/3D registration problem [7, 8, 9]. Suffered from the noisy in the input images, the former solutions require a closer initialization. In this paper, we are interested in methods in the latter category. A common disadvantage of all published solutions in this category is that they all require an explicit contour extraction as a prerequisite step, which can be quite difficult when the structure involved is complex or when the backgrounds of the X-ray images become complex.

By combining the present 3D/2D correspondence building method with our recently introduced 3D/3D surface reconstruction algorithm [16], a robust and accurate anatomy-based 3D/2D non-rigid registration algorithm has been developed. In the following section, the results of applying this algorithm to register statistical shape models of proximal femur to C-arm images of human cadaveric proximal femurs will be reported.

## 8 Experiments

We have performed experiments to evaluate the proposed method. For each case, two nearly orthogonal images are acquired. The acquired images are calibrated and registered using methods described in [2]. Results of both rigid and non-rigid 3D/2D registration are given.

**Rigid 3D/2D registration:** A plastic vertebra was used for this study. Fiducial markers were implanted for computing the ground truth of the registration transformation. Part of the registration procedure steps are shown in the top row of Figure 1 and the target registration error (TRE) calculated on those fiducial markers is 0.9 mm.

**Non-rigid 3D/2D registration:** Eleven cadaveric proximal femurs were used for this study. Part of the non-rigid 3D/2D registration steps of one example are given in the bottom row of Figure 1. The non-rigid registration accuracies were evaluated by digitizing 100 – 200 points from each surface of the cadaveric specimen and then computing the distance from those digitized points to the associated surface estimated from the images. The *in-vitro* experiments show a mean error of 1.2 mm (STD=0.2 mm), which demonstrates the robustness and accuracy of the proposed method.

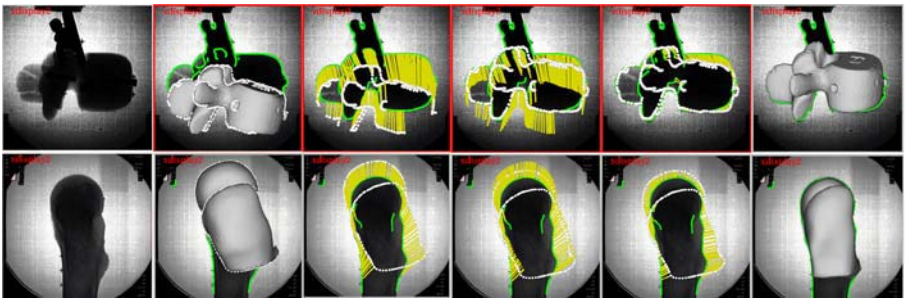


**Fig. 1.** Results of anatomy-based rigid (top row) and non-rigid (bottom row) 3D/2D registration incorporating the proposed 3D/2D correspondence building method. First column, one of the acquired images; Second column: the initial state of the 3D models (green curves: the detected edge pixels; white dots: the extracted apparent contours). Third column: after establishing 3D/2D correspondence. 3D Matched point pairs are linked with yellow line segments; Forth column: after 3D paired point matching. We apply the estimated transformation to the model and then re-calculate apparent contours; Fifth column: after re-establish 3D/2D correspondence; Sixth column: the final rigid (top) and non-rigid (bottom) registration results after a series of computations.

## 9 Conclusions

We have presented a novel 3D/2D correspondence building method and successfully applied it to both anatomy-based rigid and non-rigid 3D/2D registrations. The qualitative and quantitative results demonstrate the validity of our proposed method. In the future, we will do more studies to evaluate its robustness.

## References

1. Joskowicz L., Milgrom C., Simkin A., et al.: FRACAS: a system for computer-aided image-guided long bone fracture surgery. Comp Aid Surg, Vol. 3, No. 6, (1998) 271-288
2. Hofstetter R., Slomczykowski M., Sati M., Nolte L.-P: Fluoroscopy as an image means for computer-assisted surgical navigation. Comput Aid Surg, Vol. 4, No. 2, (1999) 65-76
3. Guéziec A., Kazanzides P., Williamson B., and Taylor R.H.: Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot. IEEE T Med Imaging, Vol. 17, No. 5, (1998) 715-728
4. Stindel E., Birard J.L., Merloz P., et al.: Bone morphing: 3D morphological data for total knee arthroplasty. Comput Aid Surg, Vol. 7, No. 3, (2002) 156-168
5. Chui H. and Rangarajan A.: A new point matching algorithm for non-rigid registration. Comput Vision and Image Under, Vol. 89, No. 2-3, (2003) 114-141
6. Zheng Y. and Doermann D.: Robust point matching for two-dimensional nonrigid shapes. Proceedings of ICCV'05, Volume 2 (2005) 1561-1566
7. Fleute M. and Lavallée S.: Nonrigid 3D/2D registration of images using a statistical model. Lecture Notes in Computer Science, Springer, Vol. 1679, (1999) 138-147.
8. le Bras A., Laporte S., Bousson V., et al.: 3D reconstruction of the proximal femur with low-dose digital stereoradiography. Comput Aid Surg, Vol. 9, Nol. 3, (2004) 51-57.

9. Laporte S., Skalli W., de Guise J.A., et al.: A biplanar reconstruction method based on 2D and 3d contours : application to the distal femur. Comput Methods Biomech Biomed Engin, Vol. 6, No. 1, (2003) 1-6
10. Hertzmann A. and Zorin D. : Illustrating smooth surfaces. In SIGGRAPH, (2000), 517-526
11. Canny J.: A computational approach to edge detection. IEEE T Pattern Anal, Vol. 8, No. 6, (1986) 679-698
12. Belongie S., Malik J., and Puzicha J.: Shape matching and object recognition using shape contexts. IEEE T Pattern Anal, Vol. 24, Nol. 24, (2002) 509-521
13. Besl P.J. and McKay N.D.: A method for registration of 3D shapes. IEEE T Pattern Anal, Vol. 14, No. 2, (1992) 239-256
14. Penney G.P., Edwards P.J., King A.P. et al.: A stochastic iterative closest point algorithm (stochastICP). Lecture Notes in Computer Science, Springer, Vol. 2208, (2001) 762-769.
15. Benameur S., Mignotte M., Parent S. et al.: 3D/2D registration and segmentation of scoliotic vertebra using statistical models. Comput Med Imag Grap, Vol. 27, (2003) 321-337.
16. Zheng G., Rajamani K.T. and Nolte L.-P.: Use of a dense surface point distribution model in a three-stage anatomical shape reconstruction from sparse information for computer-assisted orthopaedic surgery: a preliminary study. Proceedings of ACCV 2006, Part II, Lecture Notes in Computer Science, Springer, Vol. 3852, 52-60.

## Appendix: Proof of Claim 2

***Proof.*** Using proof by contradiction, let's assume that ( $I_i^s$ , $A_j^s$ ) and ( $I_m^s$ , $A_n^s$ ) are two SIN's and they also form a cross-matching and intersect at point *O* (Fig. 2).

In $\Delta OI_i^s A_n^s$ and $\Delta OA_j^s I_m^s$, applying the principle of triangle inequality, we have:

$$(| I_i^s A_j^s | + | I_m^s A_n^s |) > (| I_i^s A_n^s | + | I_m^s A_j^s |) \tag{3}$$

Additionally, as ( $I_i^s$ , $A_j^s$ ) and ( $I_m^s$ , $A_n^s$ ) are two SIN's, we have $| I_i^s A_j^s | \leq | I_i^s A_n^s |$, and $| I_m^s A_n^s | \leq | I_m^s A_j^s |$. Combining them together, we have:

$$(| I_i^s A_j^s | + | I_m^s A_n^s |) \leq (| I_i^s A_n^s | + | I_m^s A_j^s |) \tag{4}$$

Eq. (3) and (4) are contradictory to each other, i.e., ( $I_i^s$ , $A_j^s$ ) and ( $I_m^s$ , $A_n^s$ ) can not be two SIN's and also form a cross-matching. This proves the **claim 2**

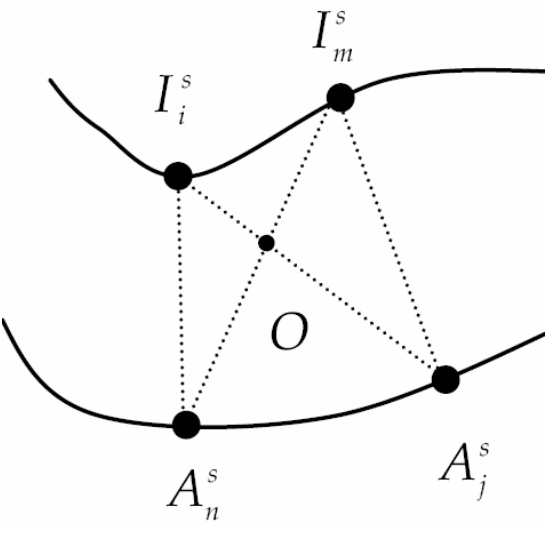



**Fig. 2.** Two SIN's that also form a cross-matching

# 2D-to-3D X-Ray Breast Image Registration

Predrag R. Bakic[1], Frederic J.P. Richard[2], and Andrew D.A. Maidment[1]

[1] University of Pennsylvania, Philadelphia, PA
{Predrag.Bakic, Andrew.Maidment}@uphs.upenn.edu
[2] Universite Paris 5, Paris, France
Frederic.Richard@math-info.univ-paris5.fr

**Abstract.** Digital breast tomosynthesis combines the advantages of digital mammography and 3D breast imaging. To facilitate the comparison of new tomosynthesis images with previous mammographic exams of the same woman, there is a need for a method to register a mammogram with tomosynthetic images of the same breast; this is the focus of our paper. We have chosen to approach this multimodality registration problem by registering a mammogram with individual tomosynthesis source projection images. In this paper, we analyzed the results of registering an MLO mammogram to nine tomosynthesis source projection images of the same breast. On average, we were able to compensate 90 percent of the per-pixel intensity differences that existed between the two images before registration.

## 1 Background

Early breast cancer detection requires identification of subtle pathological changes over time, and is often performed by comparing images from previous years. Projection mammography is considered the preferred screening modality for early breast cancer detection. However, diagnostic breast imaging is a multimodality task. Breast ultrasound is used for distinguishing cysts from solid lesions. Breast magnetic resonance imaging (MRI) offers functional information.

Recent research efforts have focused on developing 3D x-ray breast imaging modalities. Several modalities have been developed, including stereomammography, breast tomosynthesis, and breast computed tomography (CT) [1-3]. These modalities combine the advantages of mammography and 3D image visualization. The recent development of contrast-enhanced breast tomosynthesis may additionally provide functional information [4]. Of the proposed 3D x-ray modalities, breast tomosynthesis is the most likely to replace mammography as a screening procedure, chiefly because the acquisition geometry is nearly identical to mammography. In current implementations of tomosynthesis, between nine and 48 source projection images are acquired of the compressed breast as the position of the x-ray focus is altered. The total dose used is comparable to the dose needed for a mammographic exam. The projection images are used in a limited-angle CT reconstruction to form a tomographic image set. Several reconstruction algorithms have been proposed, ranging from filtered backprojection to sophisticated iterative reconstruction techniques [1,2]. Tomosynthesis produces tomographic images of the breast in which a given anatomical plane is in focus while

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 84 – 91, 2006.

anatomical structures above and below the plane are blurred to such an extent as to be essentially removed from the image.

With the clinical introduction of tomosynthesis, it will be necessary for radiologists to compare tomosynthesis images with previous mammograms of the same women to detect subtle temporal changes in the breast. It will be also necessary to compare tomosynthesis data sets of the same patient taken at different times. The former comparison task, while of a limited lifespan, requires 2D-3D registration. The latter comparison could be approached by direct registration of the reconstructed data sets. Such a registration should take into account possible differences in reconstruction algorithms used for the two 3D data sets. Alternatively, this comparison can be approached as a 2D-3D problem, in which one registers the tomosynthesis source projection images from two exams. We can look at research in computer-aided diagnosis to support the choice of 2D-3D methods being used to process tomosynthesis images. Chan *et al.* [5] are using 3D processing methods for the detection of lesions in tomosynthesis data sets, while Nishikawa *et al.* [6] use separate processing of 2D source images.

Our current research focus is on the registration of a mammogram and individual tomosynthesis images of the same breast. In this paper we present preliminary results obtained by registering an MLO mammogram and nine individual tomosynthesis source projection images obtained from one patient.

## 2 Methods and Materials

The problem of registering mammograms and tomosynthesis images can be approached in two ways. *First*, one could try to address directly the registration of a mammogram and a set of reconstructed tomographic images. This is a true multimodality registration problem. Consider the problem of finding the position in a tomographic data set which corresponds to a lesion identified in a mammogram. In this registration schema, one would need to analyze all reconstructed tomographic planes, since each plane contains only a subset of the tissue structures which are visible in the mammogram.

*Alternatively*, one could initially perform the registration of a mammogram and one or more of the projection images; this is a 2D registration problem. Each projection image should contain basically the same tissue structures as the mammogram, with some variation in positioning, compression, and dose. This registration schema, applied in multiple projection images, would allow the lesion to be located in 3D from knowledge of the acquisition geometry.

In this paper, we focus on the registration of the medio-lateral oblique (MLO) mammogram and the tomosynthesis source projection images of the same breast. In a companion paper, we analyzed the registration of the central source projection and the MLO mammogram [7]. The central projection is acquired in essentially the same MLO breast position, but with a reduced dose. The non-central projections are acquired with the same breast positioning and compression, but with the x-ray focus in different locations.

At the Hospital of the University of Pennsylvania, tomosynthesis projection images are acquired on a Senographe 2000D (General Electric, Milwaukee, WI) which has been modified to allow independent motion of the x-ray tube head. The x-ray tube can

be reproducibly positioned at nine locations, each separated by 6.25 degrees. In the current system the collimator variably occludes the detector (*see Fig. 1*). Each breast is compressed in an MLO position. The projections are acquired at a total dose equal to the dose of two-view mammography. Tomographic images are reconstructed, in planes parallel to detector, using a filtered backprojection algorithm.

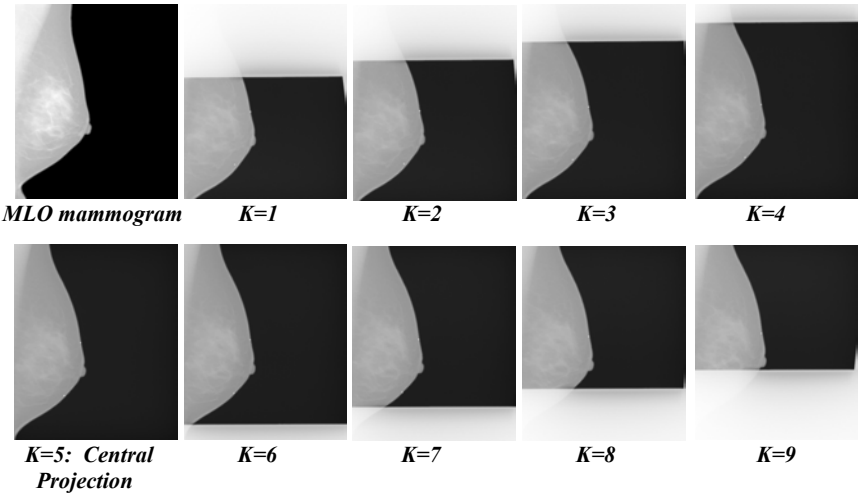



**Fig. 1.** The clinical images used for the registration of an MLO mammogram (upper left) and the tomosynthesis source projection images ($K=1,...,9$) of the same breast. The images were acquired the same day by the same technologist, with nearly the same breast positioning. The mammogram and the central projection ($K=5$) were acquired with the same geometry, but with different dose. The non-central projections ($K\neq5$) were acquired with different x-ray focus locations.

We use a non-rigid method to register the MLO mammogram and the tomosynthesis projection images of the same patient. The registration method combines intensity- and contour-based constraints to match regions of interest (ROIs) in the source and target images [8]. The registration task is formulated as the inverse problem of finding a geometric deformation that minimizes an energy function with free boundary conditions. The energy function includes three constraints designed (i) to prevent ill-posed solutions by regularization, (ii) to compensate for linear variations in image intensities, and (iii) to correct the initial mapping of the ROI in target image onto the corresponding ROI in source image. Before the registration, the ROIs in the source and target images were identified as the breast regions without the pectoral muscle. The pectoral muscle area was identified as the region above the line defined by two manually selected points on the muscle contour. In addition, the region occluded by the collimator were manually identified in each tomosynthesis projection image, and replaced by pixels of zero intensity; the same region in the mammogram was also replaced by pixels of zero intensity.

In this study, we registered the two images by deforming the mammogram to match the individual tomosynthesis projections of the same breast. The non-rigid registration method was performed in two steps. First, an initial registration was performed, based on the contour matching only. This initial step is followed by the corrections of the differences in the pixel intensity distribution between the target and source images. Detailed description of the registration method is given in our previous publications [8]. In an evaluation using synthetic images generated with a software breast model [9], an average displacement error of 1.6 mm was obtained for mammograms with compression differences of up to 3 cm. [10]. This is acceptable, as we have observed that the compression difference between mammography and tomosynthesis is approximately 1 cm.

To date, 51 clinical breast tomosynthesis exams have been performed as a part of an IRB approved clinical study in our institution. After providing informed consent, each patient in the study also received digital or film-screen mammography on the same day. As a result, there are only a few, specific variations that can exist between the images (*see Fig. 1*). This is of importance for initial testing of the registration methods because no temporal changes in the breast tissue will have occurred.

We evaluated the registration results by calculating the percentage of corrected differences, PCQD, defined as:

$$PCQD = [\Sigma_{ij}(\Delta^2_{ij})^{PRE} - \Sigma_{ij}(\Delta^2_{ij})^{POST}] / \Sigma_{ij}(\Delta^2_{ij})^{PRE} \times 100\% \quad (1)$$

where $(\Delta^2_{ij})^{PRE}$ and $(\Delta^2_{ij})^{POST}$ represent the quadratic differences between the intensities of the pixels at position $(i,j)$, before and after registration, respectively. $(\Delta^2_{ij})^{P} = [M(i,j)^{P} - T_K(i,j)]^2$, $(K=1,...,9)$, where $M(i,j)^{P}$ represents the intensity of the pixel at position $(i,j)$ in the mammogram, before $(P=PRE)$ or after $(P=POST)$ registration, and $T_K(i,j)$ represents the intensity of the pixel at position $(i,j)$ in the $K$th tomosynthesis projection. The higher PCQD values indicate the better registration performance. We also compared the root-mean-square (RMS) difference between the mammogram and the projection image, computed before and after registration:

$$RMS\ Image\ Difference = [\Sigma_{ij}(\Delta^2_{ij})^{P}]^{1/2},\ (P=PRE,\ POST). \quad (2)$$

## 3 Results

Fig. 1 shows the mammogram and the nine tomosynthesis source projection images acquired from the analyzed case. Fig. 2 focuses on the registration of the mammogram (upper left) to one of the tomosynthesis source projection images (upper right). The selected projection image is labeled $K=2$ in Fig. 1. The registration result (middle image) is shown in the form of a mammographic image non-rigidly deformed to match the tomosynthesis source projection. We have evaluated the registration performance using the difference images shown in the lower row of Fig. 2. The difference images were computed before (lower right) and after (lower left) registration.

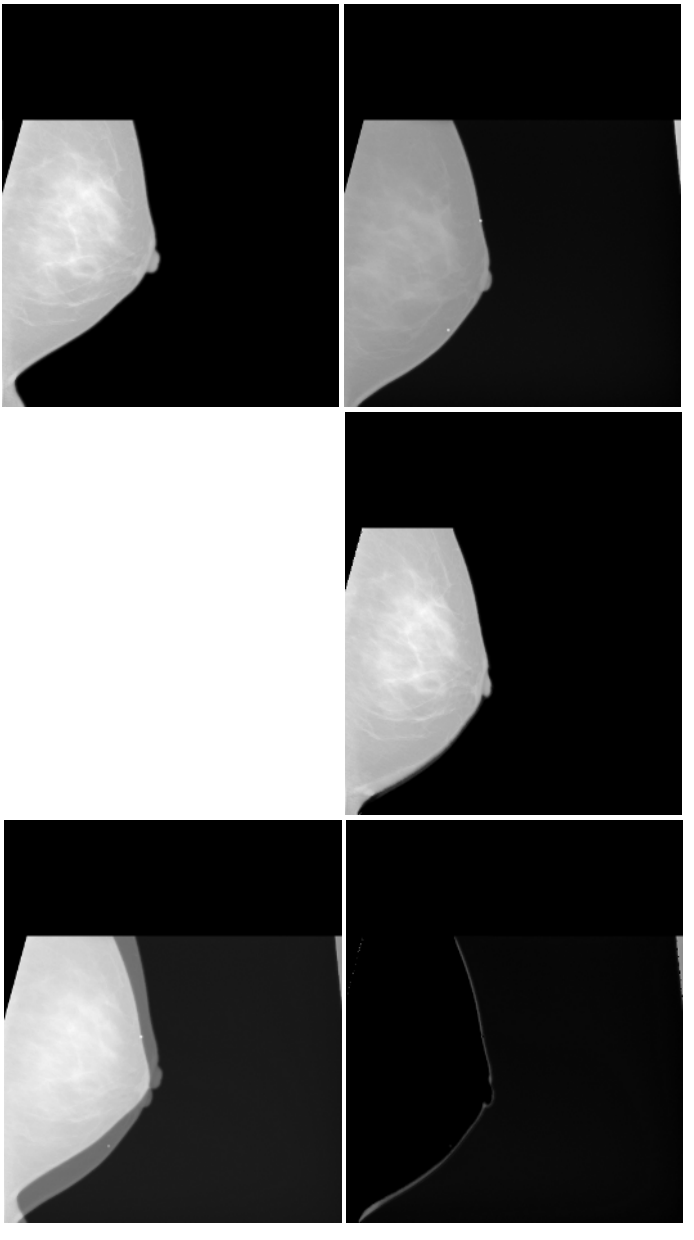
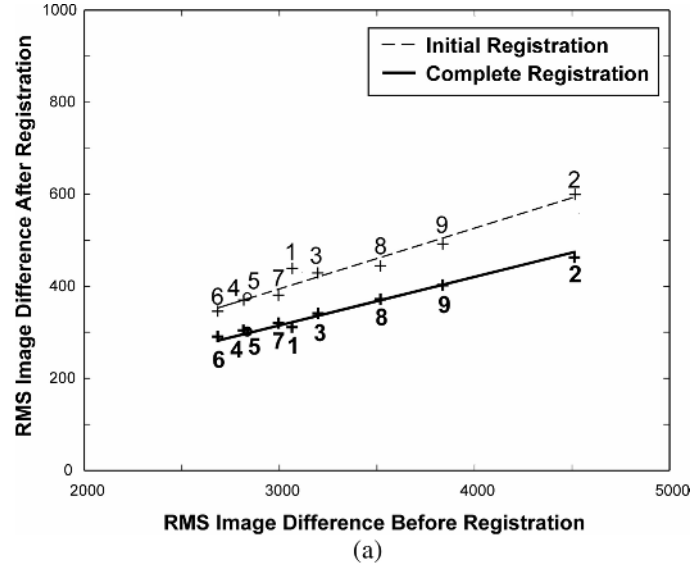
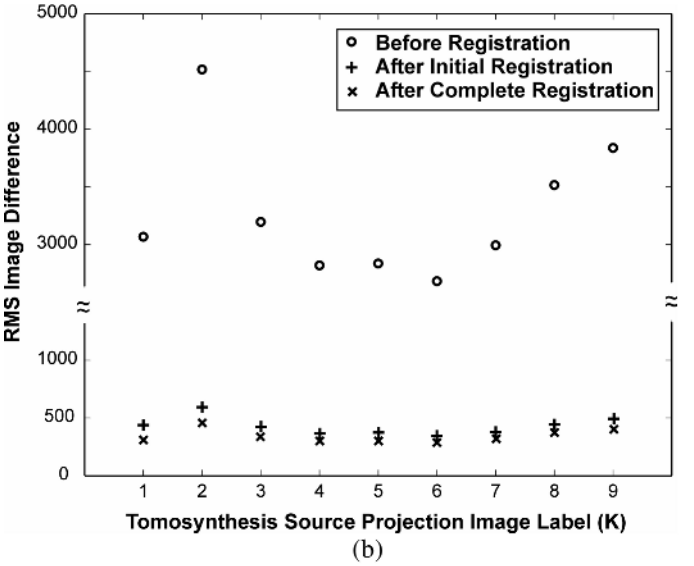


**Fig. 2.** Illustration of the registration of a mammogram and tomosynthesis source projection image of the same breast. The upper row shows the registration image pair: a mammogram (left) to be registered onto a tomosynthesis projection image (right); projection *K=2* (Fig. 1) was used. The registration result is shown in the middle row. The lower row shows the difference between the mammogram and the source projection, computed before (*left*) and after (*right*) registration.




(a)




(b)

**Fig. 3.** (a) RMS differences between the mammograms and individual tomosynthesis source projections of the same patient, computed before and after registration. The RMS image differences for each of the nine tomosynthesis projections are indicated by numbers 1-9; *solid* and *bold* numbers correspond to the differences computed after the initial and complete registration, respectively. The corresponding linear regressions are plotted by the dashed and bold lines, respectively. (b) RMS image differences as a function of x-ray focus location corresponding to different source projections.

We have registered the mammogram to all nine tomosynthesis source images. Using the difference images we computed the PCQD measure of the registration performance, defined in Eq. (1), after the initial registration and after the complete registration. The average values of PCQD ± one standard deviation, were equal to 58±4% and 90±4%, after the initial and complete registrations, respectively. Fig. 3(a) shows a plot of the RMS differences between the mammograms and central tomosynthesis projections, computed before and after non-rigid registration. The slope values of the linear regressions computed after the initial and the complete registration are equal to 0.13 and 0.10, respectively. Fig. 3(b) shows the RMS images differences values as function of the tomosynthesis source projection image label *K*.

## 4 Discussion

We have chosen to approach the registration of a mammogram and a tomosynthesis data set of the same breast, starting from the simpler problem of registering a mammogram and the individual tomosynthesis source projection images. The mammograms and tomosynthesis images were acquired on the same day by the same technologist, thus having minimal variations.

The computed average PCQD values are consistent with those computed in our study of non-rigid registration of mammograms and central tomosynthesis projections from 15 clinical breast image pairs [3]; in that study we computed the average PCQD values before and after registration of 52±20% and 94±3%, respectively.

Fig. 3 suggests that the image differences computed after the registration show relatively low dependence on the differences computed before the registration; the slope of the linear regression corresponding to the complete registration, shown in Fig. 3(a), is equal to 0.10. This result is also comparable to that obtained in our analysis of the registration of mammograms and central tomosynthesis projections [3]; the slope of the linear regression in that study was equal to 0.20.

Fig. 3(b) shows a variation in the RMS image differences computed before registration as a function of the tomosynthesis source projections (i.e. different x-ray focus locations). Ideally, assuming no changes in breast positioning, the minimum RMS image difference between the mammogram and the tomosynthesis source projection image should correspond to the central source projection (labeled *K=5*). Tomosynthesis projection images acquired with a larger angle to the central projection should result in an increased image difference. In Fig. 3(b), the minimum image difference is observed for the source projection *K=6*. The observed variation is not significant. Small changes in breast positioning between the mammography and tomosynthesis exam could cause this observation. Another possibility is that the calculation of the RMS image difference is sensitive to the variable occlusion of the detector (*see Fig. 1*). This latter issue is resolved in a new Senographe DS digital mammography machine (General Electric, Milwaukee, WI), optimized for the use in tomosynthesis, which is being installed in our department.

## 5 Conclusions

We performed a non-rigid registration of a clinical MLO mammogram with nine tomosynthesis source projection images of the same woman. Individual tomosynthesis source projection images were acquired at different positions of the x-ray tube, each separated by 6.25 degrees. The mammograms and tomosynthesis images were acquired on the same day by the same technologist, thus having minimal variations. We evaluated the registration performance by computing the percent corrected quadratic differences between the mammogram and the central tomosynthesis projection. On average we were able to compensate 90 percent of the per-pixel intensity differences that existed between the two images before the registration. In this paper, we evaluated the registration performance based on the pixel intensity differences computed from clinical images of a single patient. We are currently expanding this work to include more patients and to evaluate the registration results based on the average displacements of manually or automatically extracted fiducial points.

## References

1. Niklason, L.T., Christian, B.T., Niklason, L.E., Kopans, D.B., et al.: Digital Tomosynthesis in Breast Imaging. Radiology. 205 (1997) 399-406
2. Maidment, A.D.A., Albert, M., Conant, E.F.: Three-dimensional Imaging of Breast Calcifications. In Proc. SPIE 3240 (1998) 200-208
3. Bakic, P., Richard, F.J.P., Maidment, A.D.A.: Registration of Mammograms and Breast Tomosynthesis Images. In Proc. 8th Int. Workshop Digital Mammography, Manchester, UK (2006) in press
4. Boone, J.M., Kwan, A.L.C., Nelson, T.R., Shah, N., et al. : Performance Assessment of a Pendant-Geometry CT Scanner for Breast Cancer Detection. In Proc. SPIE 5745. (2006) 319-323
5. Chan, H.P., Wei, J., Sahiner, B., Rafferty, E.A., Wu, T., Roubidoux, M.A., Moore, R.H., Kopans, D.B., Hadjiiski, L.M., Helvie, M.A.: Computer-aided Detection System for Breast Masses on Digital Tomosynthesis Mammograms: Preliminary Experience. Radiology 237 (2005) 1075-1080
6. Reiser, I., Nishikawa, R.M., Giger, M.L., Wu, T., Rafferty, E.A., Moore, R., Kopans, D.B.: Computerized Mass Detection for Digital Breast Tomosynthesis Directly from the Projection Images. Med. Phys. 33 (2006) 482-491
7. Carton, A.-K., Li, J.J., Albert, M., Chen, S.C., et al.: Quantification for Contrast-enhanced Digital Breast Tomosynthesis. In Proc. SPIE 6142. (2006)
8. Richard, F.J.P., Cohen, L.: Non-rigid Image Registration with Free Boundary Constraints: Application to Mammography. Comp Vis Image Understanding. 89 (2003) 166-1967.
9. Bakic, P.R., Albert, M., Brzakovic, D., Maidment, A.D.A.: Mammogram synthesis using a 3D simulation. I. Breast tissue model and image acquisition simulation. Med. Phys. 29 (2002) 2131-2139
10. Richard, F.J.P., Bakic, P.R., Maidment, A.D.A.: Mammogram Registration: a Phantom-Based Evaluation of Mammographic Compression Effects. IEEE Trans Med Imag 25 (2006) 188-197

# Variational Image Registration with Local Properties

Sven Kabus[1,2], Astrid Franz[2], and Bernd Fischer[1]

[1] Institute of Mathematics, University of Lübeck, 23560 Lübeck, Germany
[2] Philips Research Laboratories, 22335 Hamburg, Germany
sven.kabus@philips.com

**Abstract.** In this paper we are concerned with elastic medical image registration. By spatially varying parameters, a displacement field can be reached which is adapted to local material properties. In addition, it enables the introduction of discontinuities within the displacement field inbetween different anatomical structures, like bones and soft tissue. The capability of this approach is demonstrated by various academic examples.

## 1 Introduction

Nonrigid image registration is a challenging field of growing importance in medical imaging. The task is to find a vector field of displacements such that each point in a template image can be mapped onto a corresponding point in a reference image in a 'meaningful' manner.

By the notion 'meaningful' often a type of constraint is meant which both preserves the topology and prescribes identical elastic properties throughout the image domain. However, there exist several cases where changes in topology are essential and/or where anatomical structures behave different from each other. For instance, structures which are connected in one image may be disconnected in the other image, like the brain-skull interface subject to a brain shift. Furthermore, structures may move along each other and thereby causing discontinuities, like the liver or a joint and their surrounding tissues. In addition, soft tissue is of different elasticity compared to bone structures and therefore behaves different. Also, preservation of shape or volume may be a reasonable property.

Typically, the wanted displacement is computed subject to a smoothness constraint. For example, the constraint is realized by a regularization based on the linear elastic potential of the displacement. In general, the constraint is applied globally with one global regularization parameter and – for the elastic regularizer – with elastic properties independent from the image position. Usually, such a method provides satisfactory results due to the underlying physical model. Nonetheless it fails in cases described above, since a global regularization does not allow for any local changes in topology or material properties. Therefore, in this note a 'meaningful' transformation enables changes in topology, supports local material properties, possibly approximates a shape or volume preservation

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 92–100, 2006.

and requires, to this end, a locally varying regularization. As a consequence, further a priori knowledge has to be added. This can be achieved by a segmentation of the template image only.

In the literature one can find several attempts dealing with nonrigid image registration in conjunction with spatially varying regularization or material parameters, for example the radial basis functions [1], the Bezier tensor product [2], the B-spline with subsequent filtering [3], the damped springs [4], the finite elements [5, 6, 7] or the finite differences [8] based approaches, respectively. However, these methods either do not reflect the physical behavior of the underlying material, or the registration yields a smooth transformation field, allowing for no discontinuities at all.

In [9, 10] we briefly introduced a new approach which overcomes the above mentioned shortcomings. In this note we extend the new idea and describe the method in greater detail. The following section is concerned with its mathematical formulation whereas Section 3 addresses the numerical treatment. Finally, we demonstrate its advantages by application to academic examples.

## 2 Variational Approach

Let $R, T : \Omega \to G$ denote the reference and the template image, respectively. Here, $G$ denotes a set of gray values and $\Omega \subset \mathbb{R}^d$ the $d$-dimensional image region. In addition, let a meaningful segmentation of $T$ be given. That is, a decomposition of $\Omega$ into disjoint regions $\Omega_l$ is assumed, such that $\Omega = \cup_{l=0}^{m} \Omega_l$. For convenience, let $\Omega_0$ denote the background of image $T$.

The registration aims at finding a displacement field $\boldsymbol{u} : \Omega \to \mathbb{R}^d$ such that $T_{\boldsymbol{u}} := T(\boldsymbol{id} + \boldsymbol{u})$ is similar to $R$, where $\boldsymbol{id}$ denotes the identity mapping. In mathematical terms, the similarity is described by a functional $\mathcal{D}[\boldsymbol{u}; T, R]$. $\mathcal{D}$ can be chosen as any popular distance (or similarity) measure provided its Gâteaux derivative exists. However, this note is restricted to the common sum of squared differences, $\mathcal{D}[\boldsymbol{u}; T, R] = \int_\Omega [R(\boldsymbol{x}) - T_{\boldsymbol{u}}(\boldsymbol{x})]^2 \mathrm{d}\boldsymbol{x} =: \int_\Omega L^{\mathcal{D}} \mathrm{d}\boldsymbol{x}$, which assumes monomodal images.

A registration based on a similarity measure only, may yield a deformed template image which perfectly matches the reference image as long as all gray values are present in both images. However, the problem is ill-posed and the underlying deformation does in general not make sense in a physical context. Therefore, an additional smoothness constraint (or regularizer) is considered which can be chosen to model the application specific physical properties. Also, it may be interpreted as a penalizer. In this note we investigate a regularizer based on the popular linear elastic potential which is in addition equipped with spatially varying parameters (the so-called *variable elastic regularizer*),

$$\mathcal{S}[\boldsymbol{u}; \alpha, \lambda, \mu] = \int_\Omega \alpha_{\boldsymbol{u}} \left( \frac{\mu_{\boldsymbol{u}}}{4} \sum_{i,j=1}^{d} (\partial_{x_j} u_i + \partial_{x_i} u_j)^2 + \frac{\lambda_{\boldsymbol{u}}}{2} (\nabla \cdot \boldsymbol{u})^2 \right) \mathrm{d}\boldsymbol{x} =: \int_\Omega L^{\mathcal{S}} \mathrm{d}\boldsymbol{x},$$

where $\alpha_{\boldsymbol{u}}$, $\lambda_{\boldsymbol{u}}$ and $\mu_{\boldsymbol{u}}$ are defined in analogy with $T_{\boldsymbol{u}}$. For other regularizers including diffusive-, fluidal- or curvature-based approaches we refer to, e.g., [11].

In contrast to a conventional approach, where $\alpha$, $\lambda$ and $\mu$ are global constants, all the three parameters are assumed to be spatially dependent. Here, the positive weighting function $\alpha : \Omega \to \mathbb{R}^+$ describes the local influence of the regularizer. By knowing the segmentation of the template image we are now in a position to reduce the regularization of the displacement field locally and, therefore, to allow for local changes in the topology. To this end, $\alpha \ll 1$ is set in the background region $\Omega_0$, cf. [10]. The Lamé parameters $\lambda, \mu : \Omega \to \mathbb{R}^+$ are used to reflect the material properties. From a qualitative point of view, $\mu$ is inversely proportional to the elastic modulus and $\lambda/\mu$ is proportional to the incompressibility of the material. For a detailed interpretation and a comparison of values for specific anatomical structures used in the literature we refer to [7]. Again, by exploiting the segmentation of $T$, different elastic properties can be assigned to each subdomain $\Omega_l$. Thereby diverse elastic behavior of different materials, like bones and muscles, can be simulated.

Note, that $\alpha_{\boldsymbol{u}}$, $\lambda_{\boldsymbol{u}}$ and $\mu_{\boldsymbol{u}}$ depend on the displacement $\boldsymbol{u}$. This dependency is indispensable due to the fact that nonlinear registration approaches mostly employ an iterative scheme and therefore the material properties at a fixed position do change in the course of the registration. As a consequence, the parameters at an intermediate stage can be deduced from $\boldsymbol{u}$ applied to the initial setting which makes a segmentation of the reference image redundant.

By combining the similarity measure and the regularizing term, the problem is to find a displacement field $\boldsymbol{u}$ which minimizes the joint functional

$$\mathcal{J}[\boldsymbol{u}] := \mathcal{D}[\boldsymbol{u}] + \mathcal{S}[\boldsymbol{u}] = \int_\Omega L^{\mathcal{D}} \mathrm{d}\boldsymbol{x} + \int_\Omega L^{\mathcal{S}} \mathrm{d}\boldsymbol{x}. \tag{1}$$

The computation of the Gâteaux derivative of (1) yields a necessary condition for $\boldsymbol{u}^*$ being a minimizer of (1),

$$\nabla_{\boldsymbol{u}} L^{\mathcal{D}} + \nabla_{\boldsymbol{u}} L^{\mathcal{S}} - \nabla_{\nabla \boldsymbol{u}} L^{\mathcal{S}} = \boldsymbol{0}.$$

Here, $\nabla_{\boldsymbol{u}}$ refers to the gradient with respect to $(u_1, \ldots, u_d)$ whereas $\nabla_{\nabla \boldsymbol{u}}$ denotes the gradient with respect to the Jacobian of $\boldsymbol{u}$. The outcome is a system of nonlinear partial differential equations equipped with associated boundary conditions,

$$\begin{aligned} \mathcal{A}\boldsymbol{u} + \boldsymbol{g}(\boldsymbol{u}) + \boldsymbol{f}(\boldsymbol{u}) &= \boldsymbol{0} \qquad \text{on } \Omega, \\ \frac{\partial u_i}{\partial \boldsymbol{n}} &= 0 \qquad \text{on } \partial\Omega,\ i = 1, \ldots, d, \end{aligned} \tag{2}$$

where $\boldsymbol{f}(\boldsymbol{u}) := -(R - T_{\boldsymbol{u}}) \nabla T_{\boldsymbol{u}}$ results from differentiating the similarity measure and is therefore independent from the choice of a regularizer. For the variable elastic regularizer a straightforward calculation yields

$$\begin{aligned} \mathcal{A}\boldsymbol{u} &:= -\nabla \cdot [\alpha_{\boldsymbol{u}} \mu_{\boldsymbol{u}} (\nabla \boldsymbol{u} + \nabla^{\mathsf{T}} \boldsymbol{u})] - \nabla [\alpha_{\boldsymbol{u}} \lambda_{\boldsymbol{u}} \nabla \cdot \boldsymbol{u}] \quad \text{and} \\ \boldsymbol{g}(\boldsymbol{u}) &:= \frac{1}{4} \sum_{i,j=1}^{d} (\partial_{x_j} u_i + \partial_{x_i} u_j)^2 \nabla [\alpha_{\boldsymbol{u}} \mu_{\boldsymbol{u}}] + \frac{1}{2} (\nabla \cdot \boldsymbol{u})^2 \nabla [\alpha_{\boldsymbol{u}} \lambda_{\boldsymbol{u}}]. \end{aligned} \tag{3}$$

Here, we collected terms with a linear dependency on $\boldsymbol{u}$ or on its derivatives in $\mathcal{A}\boldsymbol{u}$ and those with a nonlinear dependency in $\boldsymbol{g}(\boldsymbol{u})$. This allocation will become handy in the numerical treatment. Note that $\mathcal{A}\boldsymbol{u} = \boldsymbol{g}(\boldsymbol{u}) + \boldsymbol{f}(\boldsymbol{u})$ corresponds to the Navier-Lamé equations. The boundary conditions in (2) are of Neumann type but clearly they may be chosen problem dependent.

## 3 Numerical Treatment

By introducing an artificial time variable, (2) can be linearized as

$$(\boldsymbol{id} + \tau\mathcal{A})\boldsymbol{u}^{(k+1)} = \boldsymbol{u}^{(k)} - \tau\boldsymbol{f}(\boldsymbol{u}^{(k)}) - \tau\boldsymbol{g}(\boldsymbol{u}^{(k)}), \tag{4}$$

where $\boldsymbol{u}^{(k+1)} := \boldsymbol{u}(\boldsymbol{x}, t_{k+1}) = \boldsymbol{u}(\boldsymbol{x}, t_k + \tau)$ and $\boldsymbol{u}^{(0)} \equiv \boldsymbol{0}$. Due to the allocation into $\mathcal{A}$ and $\boldsymbol{g}$, the differential operator $\boldsymbol{id} + \tau\mathcal{A}$ is linear.

The system of partial differential equations (4) can be discretized on a staggered grid using second order finite differences yielding a $d \times d$ block matrix. It turned out to be reasonable to discretize (3) without evaluating the divergence operator first. Otherwise the matrix will be non-symmetric for varying parameters. As a consequence, the discretized form of (3) requires the evaluation of $\alpha, \lambda$ and $\mu$ on interlaced grid positions. Whereas $\lambda$ and $\mu$ could be interpolated either on a full- or on a half-integer grid, the definition of $\alpha$ on a half-integer grid is crucial. For example consider two adjacent anatomical structures. A displacement independently chosen for both structures requires a reduced regularity inbetween (i.e. a thin gap of background region). By defining $\alpha$ on the full-integer grid, a separate row (column) would be needed to incorporate the reduced regularity. When coarsening the scale the same row (column) would still be needed becoming more and more dominating compared to the size of the adjacent image structures. In contrast, defining $\alpha$ on a half-integer grid does not increase the dominance of the gap and is therefore recommended for a multiscale approach. However, a minimum gap size of inter-voxel width is required on the finest image level.

For stability reasons, derivatives of $\boldsymbol{g}$ are approximated by the minmod slope technique [12].

The arising system of equations is of size $dN$ ($N$ being the total number of voxels in $\Omega$). This system has to be solved at every iteration step. The system matrix resembles the Navier-Lamé differential operator and includes the additional information given by the segmentation and local parameters. The righthand-side results from both the similarity measure and further derivative terms due to the dependency of the parameters on $\boldsymbol{u}$.

Finally, to evaluate the deformed template image $T_{\boldsymbol{u}}(\boldsymbol{x})$ and to build up the linear system of equations for the following iteration step, interpolation for $\alpha_{\boldsymbol{u}}$, $\lambda_{\boldsymbol{u}}$ and $\mu_{\boldsymbol{u}}$ is required.

From a theoretical point of view the variational approach and its numerical treatment is suitable for any dimension. However, in this note we only report on results for the more instructive 1D and 2D cases. For a practical treatment, multiresolution and multigrid techniques are advisable.

## 4 Results

The proposed method has been applied to various academic images. Note that in 1D the variable elastic regularizer simplifies to a variable diffusive regularizer, cf. [10].

**1D gap example.** In order to outline some fundamental properties of the new approach we start with a 1D image consisting of five objects (cf. Fig. 1, top left, for the template image). Each object (given by an interval with non-zero gray values) belongs to a single region $\Omega_l$, $l = 1, \ldots, 5$, which is encoded in the segmented template image by assigning an integer value to each region (center left). For the outer objects there is no change in position during transition from the template image to the reference image (top of second column). The other ones are designed, such that they do change their positions in such a way that gaps between them show up or disappear. From the segmented template image we deduce the values of the weighting function $\alpha$ (bottom left). By setting $\alpha$ small in background regions we expect a displacement function which is constant within each object and inhibits high gradients inbetween.

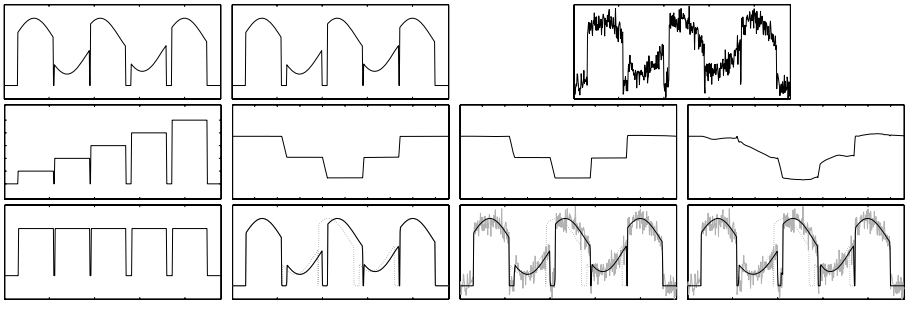


**Fig. 1.** The template image together with its segmentation and the deduced values for $\alpha$ are shown in the first column (from top to bottom). The second column displays the reference image (top), below the resulting displacement function (center) as well as the transformed template image (bottom), $T_{\boldsymbol{u}}$. In the upper right corner a modified reference image with added 10% white noise is shown. It serves for the results of the third and the fourth column where a varying $\alpha$ and a constant $\alpha$ are chosen, respectively.

As it is apparent from the second column, the variable regularizer applied with $\alpha(x) = 10$ and $\alpha(x) = 0.01$ inside and outside the objects, respectively, nicely fulfills our expectations. The displacement function (center) indicates a constant displacement within the objects with abrupt changes inbetween. Below the transformed template image is depicted. For better comparison we added the (undeformed) template image (dotted line) as well as the reference image (light gray; not visible here due to the coincidence with $T_{\boldsymbol{u}}$). Note that a similar result may be reachable when applying a constant but very small $\alpha$. However,

this problem is becoming more and more ill-posed the smaller $\alpha$ is and requires, therefore, a smaller step size $\tau$.

To test the proposed method for a more realistic setting, we modified the reference image by adding white noise with a standard deviation of 10% of the previous gray value scale (Fig. 1, top right). The template image and $\alpha$ remain unchanged. The ideal displacement field for this setting remains the same as with the unchanged reference image.

Now, the regularizer has been applied with both a varying (third column) and a constant (fourth column) weighting function. Whereas the constant choice of $\alpha = 0.03$ leads to a dissatisfying result due to the presence of noise in the reference image, a variable weighting (same as for the second column) both supports a noise-independent smooth displacement within the objects and enables for high gradients in the gap regions.

**2D rotation example.** In the second example we consider the shape-preserving feasibility of the variable elastic regularizer. To this end, a template image with a square is given. A rotation by 30° yields the reference image, cf. Fig. 2. Whereas in the first experiment all parameters are chosen constant, in the second experiment $\mu$ is multiplied by 1000 in the square region. Although, after the same number of iterations, both transformed template images almost match the reference image, the varying parameter case (cf. Fig. 2, right) is clearly preferred.

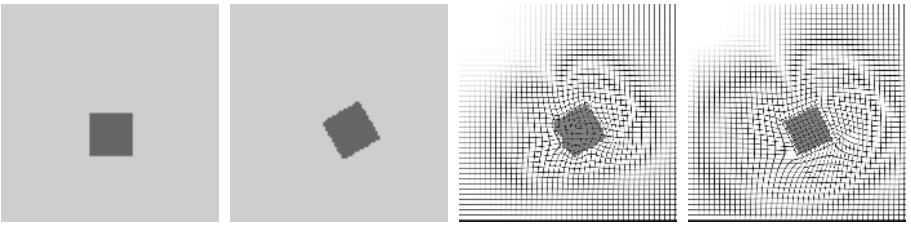


**Fig. 2.** Template (left) and reference image (center left) are displayed together with visualized displacement fields for a constant $\mu$ (center right) and a spatially varying $\mu$ (right)

**2D phantom image.** The last example considers a 2D phantom image (Fig. 3, top left) consisting of three objects: a rectangular object representing, for instance, bone structure, a square object modelling some soft tissue and in its inside a circle object taking the role of, for instance, a tumor. For the transition from the template to the reference image (Fig. 3, top right) we model a shrinking of the tissue object without affecting the bone object, which is usually a problem in registration approaches. The second problem regards the behavior of the circle object. Due to its invisibility in the reference image a conventional registration approach will tend to shrink its size in order to relate it as much as possible to a circle of zero size.

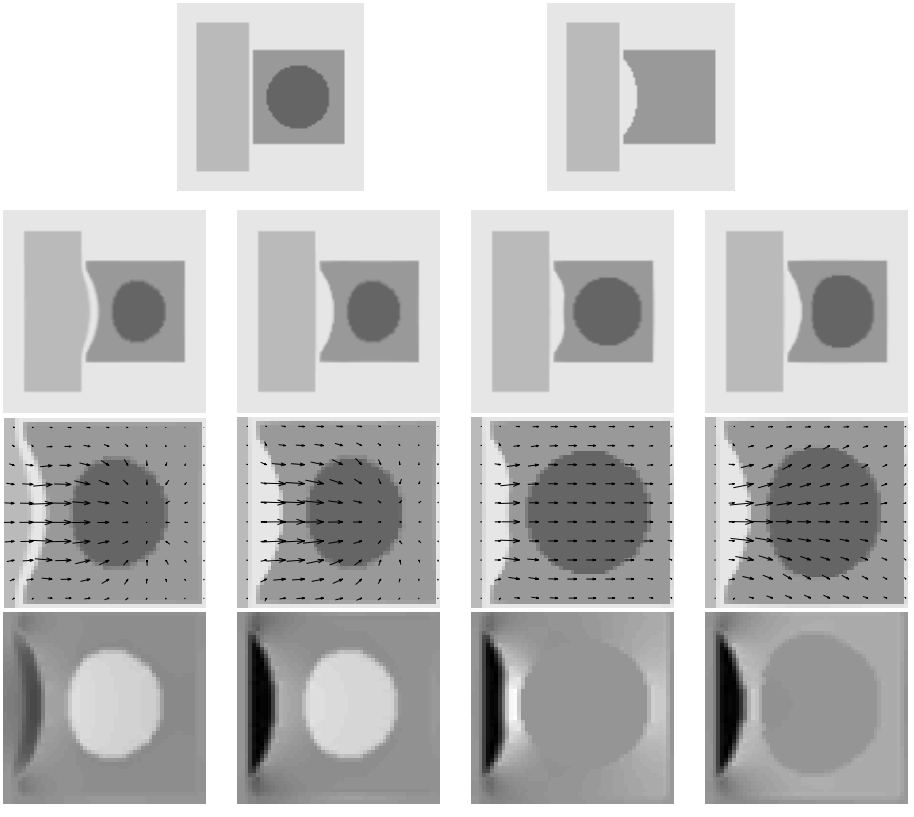


**Fig. 3.** Below the template and reference image (first row), the results from four different settings are depicted columnwise with respect to $T_{\boldsymbol{u}}$ (second row), the overlayed displacement field (third row; data are thinned out for better recognition) and the volume preservation indicator $|\mathbf{1} + \nabla \boldsymbol{u}|$ (last row), cf. text for further details

The variable elastic regularizer has been employed with four different parameter settings. For the first setting, all parameters are constant ($\alpha \equiv 0.1$, $\lambda \equiv 0.1$, $\mu \equiv 4$), cf. the first column of Fig. 3. For the remaining settings $\alpha$ is reduced locally for all background regions ($\alpha = 0.015$). In addition, for the circle object $\mu$ (cf. third column) and $\lambda$ (cf. fourth column) are multiplied by 1000, respectively.

The resulting deformation fields have been compared with respect to the deformed template image (second row in Fig. 3) and for a zoomed region around the square object with respect to the displacement field (third row) and the quantity $|\mathbf{1} + \nabla \boldsymbol{u}|$ (last row). Here, a volume preserved region (corresponding to $|\mathbf{1} + \nabla \boldsymbol{u}| = 1$) is depicted by medium gray, whereas a contracting (expanding) region appears in light gray (dark gray).

Recalling the first problem, the shrinking of the tissue object without affecting the bone object works properly whenever the weighting of the regularizer is small inbetween (second to fourth column). For the second problem several

observations can be made. With no further material knowledge the tumor object is shrinked (reduction in volume is 30%), indicated by a light gray of the circle object in the bottom row. With a large $\mu$ or $\lambda$ either a shape (and volume) preservation (third column) or an approximated volume preservation only (fourth column) can be seen. For both cases the change in volume is less than 0.3%.

## 5 Conclusion and Discussion

We have proposed an elastic potential based registration approach with displacement dependent parameters. It has been shown that this approach enables one to incorporate pre-knowledge, for instance the knowledge of anatomical structure or material properties. Whereas a proper choice of the local influence of the regularizer may lead to a discontinuous displacement field in order to model topological changes, different choices for the material parameters allow to mimic different elastic properties. Clearly, exact values for the parameters are not known in general and, usually, are guessed for in vivo situations [7].

Compared to our previous results, now, the segmentation of the template image only is sufficient. This is an important issue for time-critical tasks, like brain-shift, since an (often time-consuming) segmentation is required for the pre-operatively generated image only.

However, as a consequence from skipping the intra-operative segmentation, adjacent anatomical structures require, in order to diverge, a minimum gap of inter-voxel width inbetween. We are currently working on omitting this drawback.

## References

1. G.K. Rohde, A. Aldroubi, and B.M. Dawant. The adaptive bases algorithm for intensity-based nonrigid image registration. *IEEE Trans Med Imaging*, 22(11):1470–1479, 2003.
2. G. Soza et al. Non-linear intraoperative correction of brain shift with 1.5 T data. In *Proc. of BVM 2003*, pages 21–25, 2003.
3. M. Staring, S. Klein, and J.P.W. Pluim. Nonrigid registration with adaptive, content-based filtering of the deformation field. In *Proc. of SPIE 2005*, volume 5747, pages 212 – 221, 2005.
4. P.J. Edwards et al. A three-component deformation model for image-guided surgery. *Med Image Anal*, 2(4):355–367, 1998.
5. J. Rexilius et al. Automatic nonrigid registration for tracking brain shift during neurosurgery. In *Proc. of BVM 2002*, pages 135–138, 2002.
6. M. Ferrant et al. Registration of 3D intraoperative MR images of the brain using a finite element biomechanical model. *IEEE Trans Med Imaging*, 20(12):1384–1397, 2001.
7. A. Hagemann. *A Biomechanical Model of the Human Head with Variable Material Properties for Intraoperative Image Correction*. Logos, Berlin, 2001.

8. C. Davatzikos. Nonlinear registration of brain images using deformable models. In *Proc. of the IEEE Workshop on Math. Methods in Biomedical Image Analysis*, pages 94–103, 1996.
9. S. Kabus, A. Franz, and B. Fischer. On elastic image registration with varying material parameters. In *Proc. of BVM 2005*, pages 330–334, 2005.
10. S. Kabus, A. Franz, and B. Fischer. Variational image registration allowing for discontinuities in the displacement field. In *Proc. from the 2005 CMA Conference on Image Processing*, to appear 2006.
11. J. Modersitzki. *Numerical methods for image registration.* Oxford University Press, 2004.
12. R.J. LeVeque. *Numerical Methods for Conservation Laws.* Birkhäuser, 1992.

# Geometrical Regularization of Displacement Fields with Application to Biological Image Registration

Alain Pitiot[1] and Alexandre Guimond[2]

[1] LIDA, Brain & Body Centre, University of Nottingham, UK
[2] Siemens Molecular Imaging, Oxford, UK
alain.pitiot@nottingham.ac.uk

**Abstract.** This article tackles the registration of 2-D biological images (histological sections, autoradiographs, cryosections, etc.). The large variety of registration applications – 3D volume reconstruction, cross-dye histology gene mapping, etc. – induce an equally diverse set of requirements in terms of accuracy and robustness. In turn, these directly translate into regularization constraints on the deformation model, which should ideally be specifiable in a user-friendly fashion.

We propose an adaptive regularization approach where the rigidity constraints are informed by the registration application at hand and whose support is controlled by the geometry of the images to be registered.

We investigate the behavior of this technique and discuss its sensitivity to the rigidity parameter.

## 1 Introduction

A key component of medical image analysis, image registration essentially consists of bringing two images, acquired from the same or different modalities, into spatial alignment. This process is motivated by the assumption that more information can be extracted from an adequate merging of these images than from analyzing them independently. Its use covers a wide variety of applications, from building anatomical atlases, to longitudinal studies of tumor growth or other disease processes, through surgical planning and guidance (see [1] for a thorough overview).

More formally, given two input images, registering the source (*i.e.*, movable) image to the target (*i.e.*, fixed) image entails finding the transformation that maximizes the similarity between the transformed source image and the target one.

**Motivation.** In addition to the selected image features used to measure the adequacy of a match between two images (e.g. comparison of intensities for intensity based registration approaches, anatomically meaningful invariants such as structure boundaries for geometric techniques, etc.), another pivotal axis put

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 101–109, 2006.

forward by most registration taxonomies is the choice of a deformation model. This model acts as a regularizing scheme and embodies *a priori* information about the registration problem at hand.

Namely, in view of the characteristics of the images to be registered (less than ideal signal-to-noise ratio, tessellated structure, etc.) and of the inherently local nature of the process estimating point motions, a regularization scheme is required, both to discipline an otherwise unruly displacement field and to enforce application-dependent (i.e. user defined) constraints. In this latter respect, the selection of a suitable deformation model is crucially informed by the objectives of the medical application. For instance, in a tumor tracking system, the registration accuracy truly matters only in the region of interest defined around the identified tumor whereas the registration must be accurate everywhere in the image when building an anatomical atlas. Also of pivotal importance is the nature of the registration process. Typically, as pointed out in [2], actual anatomical deformations should not be accounted for in the same manner as variations induced by the intrinsic characteristics of the imaging modalities. For instance, the actual physical displacements generated by the growth of a tumor are different in nature from the artificial chemical shifts observed in MR, which,though only an artefact of measurement, induce an actual displacement in the image.

Essentially, we can identify two ways of controlling the rigidity of the deformation model, either directly by constraining the number of degrees of freedom of the selected transformation, or indirectly via regularization of the displacement field. We focus here on regularization approaches as they are more versatile in nature (i.e. a larger variety of transformations can be modeled within the same framework).

Among many others, we find, on the local end of the regularization spectrum, the optical flow method of Lucas and Kanadé [3] which computes a weighted average displacement in a small window centered around each pixel. In direct line with optical flow, the Demons algorithm [4] uses Gaussian filtering over a specifiable neighborhood. Noting that linear elastic models do not necessarily preserve image topology for large displacement, Christensen *et al.* [5] proposed a physics inspired viscous fluid model to enforce topological constraints within a partial differential equation framework. In Pitiot *et al.* [6], histological sections were automatically segmented into smaller components which were then registered independently, thereby implementing an anatomically guided regional regularization. On the other end of the spectrum, Barber [7] estimated a global affine transformation in a least-square sense from a correspondence field computed at each pixel to register pelvic scintillographic images. In Ourselin *et al.* [8], the use of M-estimators helped recover a global rigid or affine transformation in a robust fashion and reconstruct 3-D histological volumes. M-estimators were also selected by Hellier *et al.* [9] to deal with the inherent heterogeneity between relevant and inconsistent data on the one hand and neighborhoods with a smooth or discontinuous expected field on the other hand. They used an adaptive multri-grid multi-resolution system to register T1-weighted MRIs. Similar in spirit to our approach, Feldmard *et al.* [10] estimate a local rigid or affine

transformation in a sliding sphere to regularize the displacement field computed between two surfaces in 3-D.

Clearly, both the shape (in particular the size) of the neighborhood over which the regularizing process is applied and the nature of this process should vary as a function of the application. A common requirement of histological section registration problems is that of adequately matching the main anatomical features without inducing unsightly and biologically improbable tissue distortions. Still, as argued above, the desired rigidity of the registration process depends on the envisioned application. On the one hand, when reconstructing a histological volume, the overall transformation must remain sufficiently rigid not to induce anatomically spurious alignments in the reconstructed volume. Indeed, these would not be compatible with the anatomical reference, usually an MR image, to which it will be subsequently registered. On the other hand, when registering slices which underwent different histological treatments to reveal different genetic characteristics, the transformation space should be sufficiently flexible to allow for the resolution of local variations in fine geometrical details.

**Adaptable Rigidity.** Classical techniques usually offer only limited, and often indirect, control over the characteristics of this regularization process. More over, little if any information intrinsic to the images is taken into account. For instance, most piecewise techniques will subdivide the input images into rectangular areas which makes little anatomical sense.

In this article, we propose a regularization approach whose rigidity is informed by the registration application at hand and whose shape is controlled by the geometry (or topology) of the images to be registered. Such approach enables us to tackle a variety of registration tasks with the same generic framework.

We detail the regularization (and registration) methods in Section 2 and discuss some histological registration results in Section 3, along with the sensitivity of our algorithm to regularization parameters.

## 2 Method

To illustrate the efficacy of the proposed regularization approach, we implemented it within a classical nonlinear registration framework where B-splines were used to model the transformation space and the displacement field to be regularized was estimated with a block-matching algorithm.

This registration framework consists of a three phase, iterative process. First, a block-matching algorithm estimates a displacement field between the source and target image (Section 2.1). The field is then regularized with our method (Section 2.2). Finally, the regularized field is fed to the B-spline coefficients and the source image is resampled. This process is iterated a number of times, usually between 2 and 10 times depending on the required amount of deformation. Note that as a pre-processing step, we first globally affinely register the input images (see Section 3).

### 2.1 Block-Matching Field

We associate with the source image, $I_S$, a rectangular lattice, $L_S$, whose sites correspond to pixels in $I_S$. We consider a sparse regular site distribution, usually one site every 10 pixels, and discard sites which lay on the background as they do not correspond to actual tissue (a simple thresholding algorithm proved sufficient for histological data).

In a nutshell, for each site $s$ in $L_S$ (around which we define a rectangular neighborhood of pixels, $b_S^s$, called a "block"), the block-matching algorithm determines the block $b_T^{t_{max}}$ of the target image in an exploration neighborhood $N(s)$ centered on $s$ which maximizes a given similarity metric inside that neighborhood: $t_{max} = \arg\max_{t \in N(s)} similarity(b_S^s, b_T^t)$. The displacement at $s$ is then given by: $d(s) = centroid(b_T^{t_{max}}) - centroid(b_S^s)$ (see [6] for details).

The resulting displacement field is essentially determined by three parameters: the size of the blocks, the similarity metric and the size of the exploration neighborhood in $L_R$. The chosen similarity metric and the size of the blocks must fit the expected relationship between the intensity distributions of blocks in the source and target images, and the scale of the features of interest within those blocks respectively (see [11]). We chose the constrained correlation coefficient as a similarity metric since it offers a more appropriate block-matching computation than the conventional version [6]. The size of the exploration neighborhood is linked to the expected magnitude of the residual displacements after global alignment. It conditions the extent to which our registration algorithm can recover large deformations: the further apart corresponding components are, the larger the size of the neighborhood must be.

### 2.2 Rigidity Adaptable Regularization

As argued in the introduction section, the local nature of the pixel motion computation and the characteristics of the input images both contribute to a noisy estimated displacement field which requires subsequent smoothing. We propose to cut the regularization neighborhood to fit the geometry of the images, which

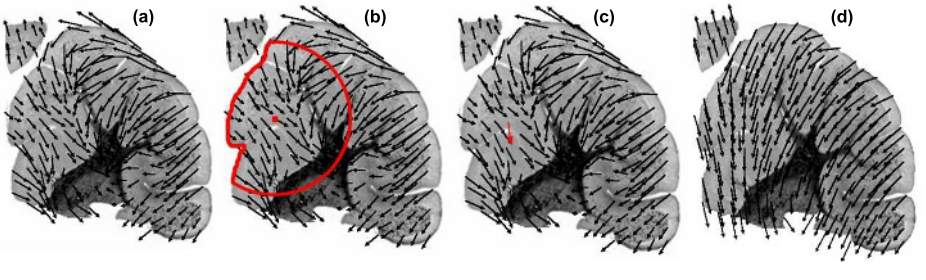



**Fig. 1.** Illustration of the rigidity adaptable regularization: (a) input noisy displacement field; (b) geometrical neighborhood around selected site; (c) estimated affine transformation applied to selected site; (d) regularized displacement field

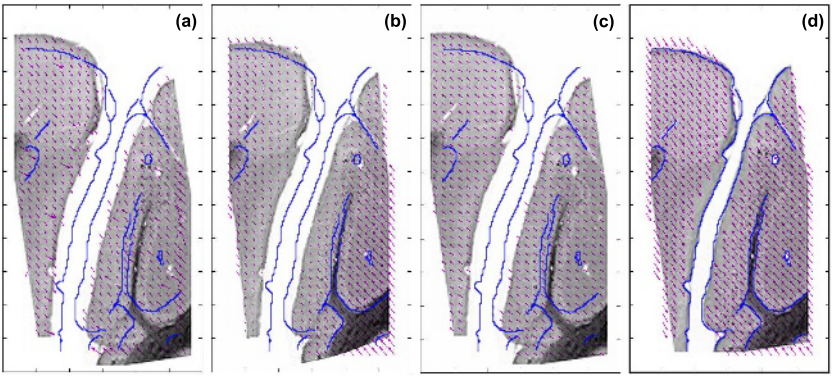



**Fig. 2.** Geometrical vs. non-geometrical regularization: (a) input source image with edges of target image and block-matching field superimposed (vectors are rescaled); (b, c & d) various displacement field regularized over geometrical or non-geometrical neighborhoods and corresponding registered source image (see text for details)

also enables us to craft a regularization parameter with more intuitive anatomical semantics.

The geometry of the source image is exploited by distinguishing between pixels belonging to actual tissues and background ones (which correspond to the glass slide supporting the tissues for histological sections, the embedding medium for cryosection, etc.). Given a site $s$ in $L_S$, we build the list of sites at distance $R$ from $s$ such that the line segment that links them does not intersect the background of the source image. That is, we only include topological neighbors. A robust least square regression algorithm (Least Trimmed Squares) is then used to estimate a rigid or affine transformation $T$ from the displacement vectors associated with the neighbor sites. At a glance, a LTS estimator differs from a conventional least square method by iteratively minimizing the sum of the $h$ smallest squared residuals, to reduce the influence of outliers [12]. We set $h$ at 70% of the number of residuals. The original displacement vector at $s$ is then replaced by $T(s) - s$. This process is repeated for every site in $L_S$ (see Figure 1).

By varying the regularization radius $R$ between 0 and $+\infty$, we control the rigidity of the regularized displacement field, and hence, that of the transformation. A null radius corresponds to a fully flexible transformation, while $R = +\infty$ yields a globally rigid transformation (or a set of rigid transformations when the image consists of multiple connected or semi-connected components, see Section 3). Furthermore, given that the topological neighborhoods associated with topologically adjacent sites are very similar, the estimated transformations will also be very similar, ensuring smooth transitions throughout the image.

Figure 2 illustrates this technique on the regularization of a displacement field computed on either side of a gap between the cerebellum and the cerebrum in a small piece of a calbindin-stained mouse brain sagittal slice. The source image was obtained from the target image by applying a left and a right translation on either side of the gap. When both sides of the gap were taken into account, local

affine estimation incorrectly yielded a compression (b). A rigid (instead of affine) estimation would have given a global left or right translation depending on which side contained the largest number of blocks, as did a Gaussian filtering approach with large standard deviation (c). However when both sides were considered independently (d), the estimated transformations were the correct left and right translations.

## 3 Results and Discussion

We tested our regularization technique on a series of 10 calbindin-stained histological sections of the left hemisphere of a transgenic mouse brain ($600 \times 400$ pixels downsampled from $6000 \times 4000$). Figure 3(a & b) shows the source (#6) and target (#7) sections Note that in addition to the 0.7mm gap in between slices which made for substantial anatomical differences (for instance, the olfactory bulb suddenly appears in the target section), tears also induced large pictorial discrepancies.

Registration consisted of a global affine registration ($R = +\infty$) followed by 3 iterations (blocks: $20 \times 20$ pixels, exploration neighborhood: $80 \times 80$ pixels,

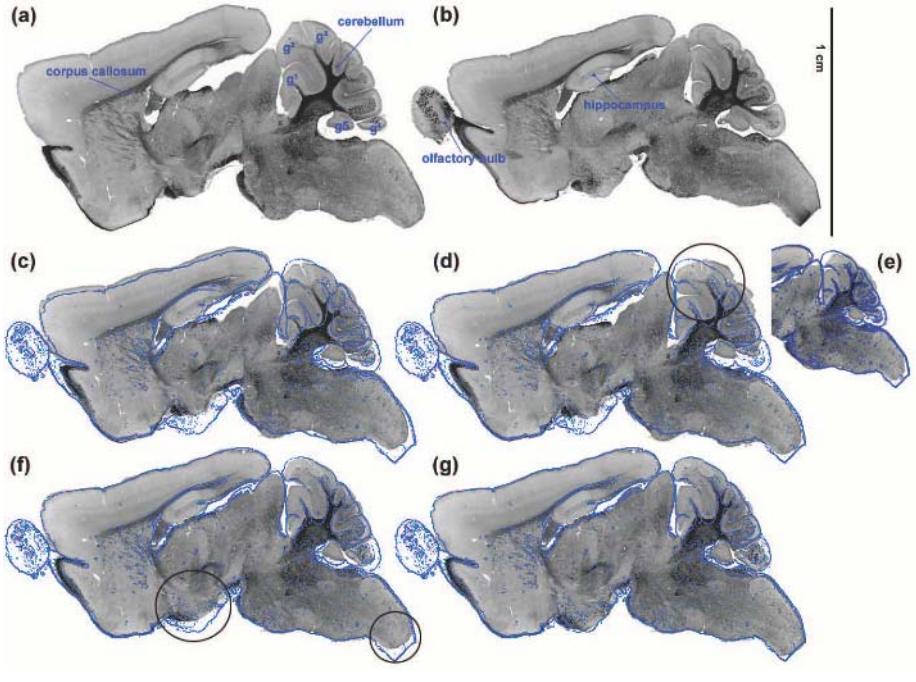



**Fig. 3.** Geometrical regularization results: (a) source image; (b) target image; registered source image with superimposed target edges: (c) global affine registration, $R = +\infty$; (d) local rigid registration, $R = 100$; (e) local affine registration, $R = 100$; (f) local affine registration, $R = 20$; (g) local affine registration, $R = 5$

B-spline control points every 10 pixels) with the specified $R$. By varying both $R$ and the nature of the locally estimated transformation (affine or rigid), we can control the characteristics of the registration. In Figure 3(d), locally rigid transformations were estimated over a large topological neighborhood ($R = 100$ pixels). Note in particular how the cerebellum gyri of the source section (identified as $g^1 - g^5$ in (a) ) were correctly repositioned. Yet, they did not quite fit inside their counterparts in target section, which was to be expected since their actual shapes and sizes had changed in between. A better match was obtained in (e) where affine transformations were estimated instead of rigid ones. This however came at the price of a large distortion in $g^3$ (see black circle in (d) ). An even more accurate correspondence was obtained in (f) (local affine transformations, $R = 20$ pixels): the major structures are perfectly aligned and so are most of the cerebellum gyri. The remaining discrepancies (gyri $g^4$, $g^5$ and circled regions in (f) ) again represent actual anatomical differences. Even though a more flexible transformation ($R = 5$ pixels in (g) ) virtually eliminates them, it also induces unpleasant distortions as the tissues are forced to stretch.

**Choosing *R*.** As illustrated above, the choice of $R$ is pivotal. We submit that it should be conditioned both by the application at hand and by the characteristics of the input images. For instance, if a 3-D volume were to be reconstructed from the calbindin sections, a large value of $R$ should be selected ($R = 100$ with affine or rigid for instance) so as not to induce in the reconstructed volume the artificial distorsions we observed in Figure 3(g). Conversely, in a genetic study interested in the correlation between the expression of two different proteins or markers (typically even and odd sections will then be treated differently to reveal different proteins or genetic markers), a small $R$ should be preffered as we would like the gyri to match as best as possible.

In our experience, a good choice for $R$ is of the order of the size of the smallest structure we wish to match across images without altering its shape. Indeed, the use of a robust estimator (LTS) then ensures a relatively constant transformation to be computed inside this structure if the cut-off is sufficiently high (close to 50%).

Note that often the input images to be registered are not homogeneous in the shape or in the distribution of the structures they contain. For instance, the calbindin sections include both a highly convoluted cerebellum and a less complex cerebrum. Even better registration results can then be achieved by adapting the selected $R$ to the local geometry in the image. Preliminary results indicate that a good strategy consists in modulating the user-supplied $R$ value by an edgeness measure (gradient magnitude) in a neighborhood the size of $R$. Furthermore, the main edges can also be used to distinguish between anatomically significant neighborhoods in addition to background pixels.

## 4 Conclusion

We have presented a regularization technique which estimates a rigid or affine transformation over a neighborhood whose size is specified by the user and whose

shape depends on the geometry of the input images. This offers a less arbitrary control over the rigidity of the registration process. As opposed to standard piecewise techniques which impose a grid-like structure on the transformation space (i.e. the piecewise areas are rectangular), this approach defines more anatomically meaningful regions in following the actual contours of the image, while eliminating the need for pre-segmentation [6]. The use of a sliding window also induces natural soft transitions between regions, except in areas where we expect a discontinuous gap (i.e. across gyri). This enables the computation of arbitrarily complex deformation fields while keeping their flexibility under control (i.e., the fields are only as flexible as they need to be), very much in the fashion of a spatially adaptive fluid approach but with arbitrary shaped regions. Obviously, a multi-resolution scheme would further improve the robustness of our method by handling large displacements, at low resolutions, in a locally rigid fashion.

As mentioned throughout the paper, the choice of $R$, though it should not be fully automatized as it must depend on the application, could probably benefit from *a priori* medical expertise. We are currently developing new heuristics to tackle this issue, in addition to the size-of-smallest-structure-of-interest one. Along similar lines, a registration algorithm implementing our regularization technique should implement an application-dependent stopping criterion as well (if the registration process is iterative). Finally, we are also investigating automated and semi-automated ways to vary the regularization radius $R$ as a function of the spatial position in the images to accommodate particularly damaged sections and missing parts.

## Acknowledgements

The authors would like to thank A. MacKenzy-Graham from the LONI laboratory for providing the mouse brain sections.

## References

1. Maintz, J.B.A., Viergever, M.A.: A Survey of Medical Image Registration. Medical Image Analysis **2** (1998) 1–36
2. Cachier, P., Bardinet, E., Dormont, D., Pennec, X., Ayache, N.: Iconic Feature Based Nonrigid Registration: The PASHA Algorithm. CVIU — Special Issue on Nonrigid Registration **89** (2003) 272–298
3. Lucas, B., Kanadé, D.: An Iterative Image Registration Technique with an Application to Stereo Vision. In: Proc. of DARPA Image Understanding Workshop. (1981) 121–130
4. Thirion, J.P.: Image Matching as a Diffusion Process: an Analogy with Maxwell's Demons. Medical Image Analysis **2** (1998) 243–260
5. Christensen, G., Joshi, S., Miller, M.: Volumetric Transformation of Brain Anatomy. IEEE Transactions on Medical Imaging **16** (1997) 864–877
6. Pitiot, A., Bardinet, E., Thompson, P., Malandain, G.: Piecewise Affine Registration of Biological Images for Volume Reconstruction. Medical Image Analysis (2005) in press.

7. Barber, D.: Registration of Low Resolution Medical Images. Physics in Medical Biology **37** (1992) 1485–1498
8. Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N.: Reconstructing a 3D Structure from Serial Histological Sections. Image and Vision Computing **19** (2001) 25–31
9. Hellier, P., Barillot, C., Mémin, E., Pérez, P.: Hierarhical Estimation of a Dense Deformation Field for 3-D RobustRegistration. IEEE Transactions on Medical Images **20** (2001) 388–402
10. Feldmar, J., Ayache, N.: Rigid, Affine and Locally Affine Registration of Free-Form Surfaces. The International Journal of Computer Vision **18** (1996)
11. Dengler, J.: Estimation of Discontinuous Displacement Vector Fields with the Minimum Description Length Criterion. In: Proc. of CVPR). (1991) 276–282
12. Rousseeuw, P.: Least Median of Squares Regression. Journal of the American Statistical Association **79** (1984) 871–880

# Myocardial Deformation Recovery Using a 3D Biventricular Incompressible Model

Arnaud Bistoquet[1], W. James Parks[2], and Oskar Škrinjar[1,3]

[1] School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[2] Department of Pediatrics, Sibley Heart Center Cardiology, Children's Healthcare of Atlanta, Emory University School of Medicine, Atlanta, GA, USA
[3] Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, USA

**Abstract.** This paper presents a nonrigid image registration method for cardiac deformation recovery from 3D MR image sequences. The main contribution of this work is that the method is mathematically guaranteed to generate incompressible deformations. This is a desirable property since the myocardium has been shown to be close to incompressible. The method is based on an incompressible deformable model that can include all four cardiac chambers and has a relatively small number of parameters. The myocardium needs to be segmented in an initial frame after which the method automatically determines the tissue deformation everywhere in the myocardium throughout the cardiac cycle. The method has been tested with four 3D cardiac MR image sequences for the left and right ventricles and it has been evaluated against manual segmentation. The volume agreement between the model and the manual segmentation exceeds 90% and the distance between the model and the manually generated endocardial and epicardial surface is 1.65$mm$ on average.

## 1 Introduction

The motion of the heart wall plays an important role in the evaluation of the heart condition. A number of image-based cardiac deformation recovery techniques have been developed. Most of these methods use a specialized imaging modality, including tagged MRI [1, 2], diffusion tensor MRI [3], DENSE MRI [4], and phase velocity MRI [5]. These imaging modalities require extra acquisition time in addition to the time required to obtain anatomical (cine) MRI, which is always acquired. For this reason it is useful to have a cardiac deformation recovery method that uses only anatomical MRI. Some of the relevant methods [6, 7, 8, 9, 10] provide only the segmentation of the myocardium without recovering the displacement field. This prevents one from computing myocardial strains, which are critical for analysis of the cardiac function. In addition, there are methods that track only the endocardial and epicardial surfaces [11, 12] without computing the displacements within the heart wall. These methods also

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 110–119, 2006.

cannot be used to compute myocardial strains. Finally, it is desirable that the method automatically recovers the myocardial deformation in all the frames of the cardiac cycle, since manual interaction in all the frames is tedious and time consuming [13, 14].

In this paper we present a method for 3D cardiac deformation recovery from anatomical MR image sequences. The method computes the displacements everywhere within the heart wall which allows one to compute myocardial strains. To initialize the method, one needs to segment the myocardium only in an initial frame, after which the method automatically recovers the myocardial deformation throughout the cardiac cycle. A byproduct of the method is an automated segmentation of the myocardium in the entire image sequence. The method is mathematically guaranteed to generate incompressible deformations. This is a desirable property since the myocardium has been shown to be close to incompressible. The method is based on an incompressible deformable model that can include all four cardiac chambers and has a relatively small number of parameters.

We also note relevant work on automated cardiac deformation recovery from anatomical MR image sequences [15, 16, 17, 18], but these methods do not incorporate the incompressibility property of the myocardium and require more parameters than our method. An almost incompressible biomechanical model of the left ventricular wall was used for image-based cardiac deformation recovery in [19]. This method is not automated, i.e. it requires manual interaction in all the frames of the cardiac cycle. In [20], the incompressibility property is introduced as a tradeoff in the objective function, and the method can only track the endocardial and epicardial surfaces for the left ventricle (LV).

## 2 Methods

### 2.1 Incompressibility of the Myocardium

The myocardium is an almost incompressible material. Its constituents are mainly composed of water, which is almost perfectly incompressible. However, the myocardium is perfused with blood, which affects the total myocardial volume over the cardiac cycle. A few studies [21, 22, 23, 24] have been carried out to quantify the change of the myocardial volume over the cardiac cycle. The common conclusion of these efforts is that the total myocardial volume changes no more than 4% during a cardiac cycle. This means that the myocardium is not perfectly incompressible. However, this volume change is distributed in all three directions. Even in the areas with predominant orientation of the blood vessels, the myocardial tissue does not expand or contract by more than 2% in any direction. Given that the thickness of the myocardium of the left ventricle (LV) is up to $14mm$ [25] for a normal heart at end systole (ES), which is the maximal thickness of the myocardium wall during the heart cycle, the compressibility of the myocardium corresponds to a $0.28mm$ change in its thickness in the worst case. This is negligible because the pixel size of the images is around $1mm$. For this reason, one conclude that incompressibility is a physical property

that has to be exactly or at least approximately satisfied for the modeling of the myocardium. The 3D transformation model presented in this paper is directly based on this property.

### 2.2 3D Incompressible Model Transformation

The goal of the method is to recover the 3D cardiac deformation from image sequences. The user needs to manually segment the myocardium in one frame of the sequence, called the reference frame, which is usually the end diastole (ED) frame. This is the only user interaction needed. The transformation is applied to the segmented myocardium of the reference frame. For each frame of the sequence, the method computes a transformation that models the myocardial deformation. The location of the myocardium in frame $i$ is given by $\mathbf{m_i} = \mathbf{T_i}(\mathbf{m})$, where $\mathbf{T_i}$ is the transformation from the reference frame to frame $i$, $\mathbf{m} = (x, y, z)$ are the myocardium points in the reference frame, and $\mathbf{m_i} = (x_i, y_i, z_i)$ the corresponding points in frame $i$.

Our transformation model follows the geometry of the heart. Let $\mathbf{r}(\alpha, \beta)$ and $\mathbf{r_i}(\alpha, \beta)$ represent a parametrization of the midsurface of a cardiac chamber in the reference frame and frame $i$, respectively. The midsurface is defined here as the surface exactly half way between the endocardial and epicardial surfaces. Note that the two functions depend on the same parameters $\alpha$ and $\beta$. This means that points $\mathbf{r}(\alpha, \beta)$ and $\mathbf{r_i}(\alpha, \beta)$ correspond, i.e. they represent the same physical point in the two frames. Let $\mathbf{n}(\alpha, \beta)$ and $\mathbf{n_i}(\alpha, \beta)$ represent the surface normals of the midsurface in the reference frame and frame $i$, respectively. A point in the chamber wall in the reference frame that is at a distance $\gamma$ in the surface normal direction from the midsurface point $\mathbf{r}(\alpha, \beta)$ is located at:

$$\mathbf{m}(\alpha, \beta, \gamma) = \mathbf{r}(\alpha, \beta) + \gamma \mathbf{n}(\alpha, \beta). \tag{1}$$

The corresponding point in frame $i$ is at a distance $\gamma_i$ in the surface normal direction from the midsurface point $\mathbf{r_i}(\alpha, \beta)$, i.e. its location is:

$$\mathbf{m_i}(\alpha, \beta, \gamma) = \mathbf{r_i}(\alpha, \beta) + \gamma_i(\alpha, \beta, \gamma)\mathbf{n_i}(\alpha, \beta). \tag{2}$$

Control points are used to represent the midsurface in the reference frame, and their coordinates represent the parameters of the transformation model. The location of the control points are automatically determined in the segmented myocardium of the reference frame. They are located on the skeleton of the myocardium on each slice of the 3D image. A pseudo thin plate spline interpolation scheme is used to generate the midsurface of the cardiac chamber from the control points [26]. The number of control points is chosen such that the mean square difference between the generated midsurface and the skeleton of the myocardium for each slice of the image is less than half a pixel. The cardiac deformation recovery explained in Sec. 2.3 provides the locations of the control points in each frame of the cardiac sequence. By using the same interpolation scheme, we define the midsurface of the cardiac chamber in each frame of the cardiac sequence.

The goal is, for given midsurfaces in the reference frame and frame $i$, to find $\gamma_i$ such that the transformation from the reference frame to frame $i$ is incompressible. Let $dv$ represent the infinitesimal volume defined by function $\mathbf{m}$ when $\alpha$, $\beta$, and $\gamma$ are varied by infinitesimal increments $d\alpha$, $d\beta$, and $d\gamma$, respectively. Similarly, let $dv_i$ represent the infinitesimal volume defined by function $\mathbf{m_i}$ when $\alpha$, $\beta$, and $\gamma$ are varied by infinitesimal increments $d\alpha$, $d\beta$, and $d\gamma$, respectively. The transformation is incompressible if $dv = dv_i$ for any $\alpha$, $\beta$, and $\gamma$. One can obtain $dv$ by taking the triple scalar product of $\frac{\partial \mathbf{m}}{\partial \alpha} d\alpha$, $\frac{\partial \mathbf{m}}{\partial \beta} d\beta$, and $\frac{\partial \mathbf{m}}{\partial \gamma} d\gamma$, i.e.:

$$dv = \frac{\partial \mathbf{m}}{\partial \gamma} \cdot (\frac{\partial \mathbf{m}}{\partial \alpha} \times \frac{\partial \mathbf{m}}{\partial \beta}) d\alpha d\beta d\gamma.$$

Using Eq. 1, the expression for the infinitesimal volume can be reduced to:

$$dv = A(1 - 2H\gamma + K\gamma^2) d\alpha d\beta d\gamma,$$

where $A = |\frac{\partial \mathbf{r}}{\partial \alpha} \times \frac{\partial \mathbf{r}}{\partial \beta}|$, and $H$ and $K$ are the mean and Gaussian curvatures, respectively, of the midsurface in the reference frame. Since $Ad\alpha d\beta d\gamma$, $H$ and $K$ are independent of parametrization [27], the above expression holds for any parametrization. It is assumed that surface normals are oriented outward. Similarly:

$$dv_i = \frac{\partial \mathbf{m_i}}{\partial \gamma} . (\frac{\partial \mathbf{m_i}}{\partial \alpha} \times \frac{\partial \mathbf{m_i}}{\partial \beta}) d\alpha d\beta d\gamma,$$

which simplifies to:

$$dv_i = A_i(1 - 2H_i\gamma_i + K_i\gamma_i^2) \frac{\partial \gamma_i}{\partial \gamma} d\alpha d\beta d\gamma,$$

where $A_i = |\frac{\partial \mathbf{r_i}}{\partial \alpha} \times \frac{\partial \mathbf{r_i}}{\partial \beta}|$, and $H_i$ and $K_i$ are the mean and Gaussian curvatures, respectively, of the midsurface in frame $i$. The incompressibility requirement ($dv = dv_i$) leads to the following differential equation:

$$1 - 2H\gamma + K\gamma^2 = S_i(1 - 2H_i\gamma_i + K_i\gamma_i^2) \frac{\partial \gamma_i}{\partial \gamma},$$

where $S_i = \frac{A_i}{A}$. After integration, the differential equation becomes a cubic equation in $\gamma_i$:

$$\gamma - H\gamma^2 + \frac{1}{3}K\gamma^3 + C_i = S_i(\gamma_i - H_i\gamma_i + \frac{1}{3}K_i\gamma_i^3),$$

where $C_i$ depends on $\alpha$ and $\beta$ but not on $\gamma$. The transformation needs to map points on the midsurface in the reference frame ($\gamma = 0$) to points on the midsurface in frame $i$ ($\gamma_i = 0$), which implies that $C_i = 0$. The distance $\gamma_i$ is obtained by solving this cubic equation. Thus we can derive the location of the myocardium in frame $i$, $\mathbf{m_i}$ given by Eq. 2.

This transformation can easily be extended to all the heart chambers. The final transformation to recover the deformation of all the chamber walls is:

$$\mathbf{T} = \mathbf{T_{RA}} \circ \mathbf{T_{LA}} \circ \mathbf{T_{RV}} \circ \mathbf{T_{LV}}, \tag{3}$$

where $\mathbf{T_{RA}}$, $\mathbf{T_{LA}}$, $\mathbf{T_{RV}}$, and $\mathbf{T_{LV}}$, are transformations of type defined in the one chamber case, for the right atrium (RA), left atrium (LA), right ventricle (RV), and LV, respectively. This combined transformation is also incompressible since the Jacobian of $\mathbf{T}$ is the product of the corresponding Jacobians:

$$J_{\mathbf{T}} = J_{\mathbf{T_{RA}}} J_{\mathbf{T_{LA}}} J_{\mathbf{T_{RV}}} J_{\mathbf{T_{LV}}} = 1. \tag{4}$$

### 2.3 Cardiac Deformation Recovery

The cardiac deformation is recovered for each frame $i$ by searching for the locations of the control points that maximize a voxel based similarity measure. We use normalized mutual information (NMI) as the similarity measure, which has previously been successfully used for cardiac image registration [15]. The cardiac deformation is recovered from the reference frame where the myocardium is segmented. Thus, for all the other frames $i$, one needs to find the control points that maximize the objective function $O = NMI(\ I_{ref}(\mathbf{m}), I_i(\mathbf{T_i}(\mathbf{m}))\ )$, where $I_i$ is the image of frame $i$, $I_{ref}$ the image of the reference frame, $\mathbf{T_i}$ the transformation from the reference frame to frame $i$, and $\mathbf{m}$ the myocardium points in the reference frame.

We start the optimization process from the reference frame where the locations of the control points are known. Then, we advance in the forward direction of the cardiac cycle. The locations of the control points in frame $i$ are determined by the Powell's method [28], which uses as initialization the locations of the control points of the previous frame $i-1$, and minimizes the objective function $O$. In this way, we obtain the locations of the control points in each frame $i$ of the cardiac cycle.

## 3 Results

We have tested our method on four 3D MR image sequences, two normal cases and two pediatric patients with tetralogy of Fallot (TOF), for the recovery of LV and RV myocardium walls. To visualize the results, we have constructed the 3D myocardium wall surface for each frame of the cardiac sequence. Fig. 1 shows the results for the first normal case.

Visual inspection and two quantitative procedures have been performed to assess the accuracy of the method. Visual inspection of the myocardium surface provided by our transformation model, as shown in Fig. 1 for the first normal case and in Fig. 2 for all the four cases, suggests that the method was able to successfully recover the deformation of LV and RV myocardium walls.

The first quantitative evaluation procedure compares the myocardial wall provided by the transformation model with the one obtained by manual segmentation for the ES frame. To evaluate the cardiac deformation recovery results, we computed the true positive (TP) rate, the false positive (FP) rate and the volume ratio (VR). Let $V_m$ represent the myocardium voxels recovered by our transformation model and $V_s$ those corresponding to the manual segmentation.Furthermore, let $V_{ms}$ be the overlap of $V_m$ and $V_s$, and $\overline{V_{ms}}$

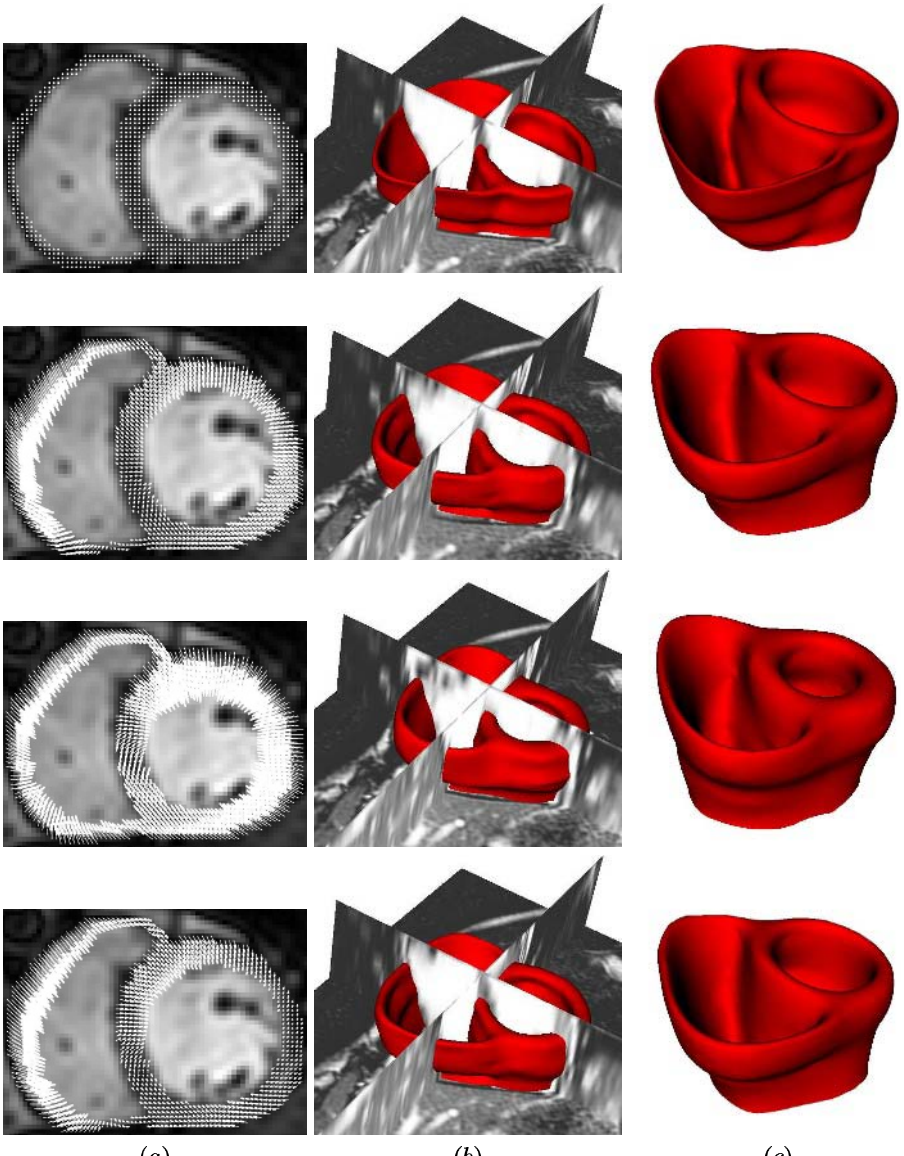



**Fig. 1.** Results for Normal 1. The rows represent different time positions of the cardiac cycle: the first row corresponds to ED, and the third one corresponds to ES (a) displacement fields from ED superimposed over the ED frame (b) recovered 3D surface superimposed with the image, (c) recovered 3D surface.

the part that is in $V_m$ but not in $V_s$. Then, $TP\ rate = \frac{V_{ms}}{V_s}$, $FP\ rate = \frac{\overline{V_{ms}}}{V_s}$, and $VR = \frac{\overline{V_m}}{V_s}$. The second quantitative procedure provides the average Euclidean distance error between the recovered LV myocardium surface and the

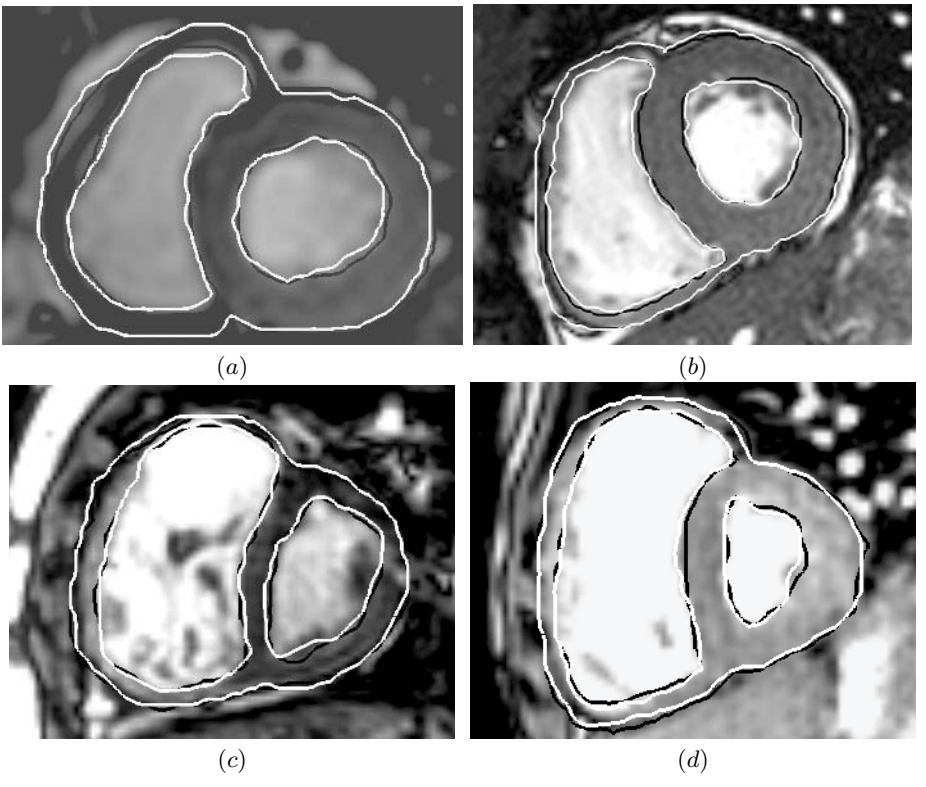



**Fig. 2.** 2D contours of the 3D myocardium wall surface for the middle slice at ES, obtained by the transformation model (black) and manual segmentation (white). (a) Normal 1, (b) Normal 2, (c) TOF 1, (d) TOF 2.

one obtained from manual segmentation. Table 1 shows for the four cases the values of these two evaluation procedures. The resulting average error of TP and FP rates is around 10%, and the volume ratio error is within 10%. The average Euclidean distance error is $1.65mm$, which is in the order of the image voxels that have a size of $1.56mm$x$1.56mm$x$8mm$. These results are similar to those reported in [15, 16, 17, 18].

## 4 Conclusion

We have developed a method for nonrigid image registration for cardiac deformation recovery from 3D MR image sequences. It is based on a 3D deformable model that guarantees incompressibility: a desirable property since myocardium has been shown to be close to incompressible. The presented method recovers the deformation throughout the myocardium, i.e. not only at the endocardium and epicardium, and it does not assume a biomechanical model, which could bias

**Table 1.** VR, TP and FP rates, and the average Euclidean distance error $d$, between the manually segmented and the recovered myocardium, are used to quantify the performance of the cardiac deformation recovery. The results are presented for the four cases ( two normal and two TOF cases) for the ES frame. The number of control points used in each case are reported.

| Case | Normal 1 | Normal 2 | TOF 1 | TOF 2 |
|---|---|---|---|---|
| **Number of control points** | 102 | 91 | 86 | 72 |
| **VR LV** [%] | 92.8 | 91.4 | 90.5 | 91.2 |
| **VR RV** [%] | 87.6 | 86.5 | 89.2 | 89.2 |
| **VR myocardium** [%] | 91.3 | 90.7 | 90.4 | 90.8 |
| **TP rate LV**[%] | 92.2 | 90.9 | 90.2 | 90.9 |
| **TP rate RV** [%] | 87.1 | 86.1 | 88.8 | 88.7 |
| **TP rate myocardium** [%] | 91.1 | 90.5 | 90.1 | 90.2 |
| **FP rate LV** [%] | 4.2 | 5.6 | 5.8 | 6.2 |
| **FP rate RV** [%] | 5.9 | 6.6 | 7.2 | 7.6 |
| **FP rate myocardium** [%] | 4.5 | 5.8 | 6.2 | 6.5 |
| **d myocardium** [$mm$] | 1.58 | 1.68 | 1.66 | 1.71 |

the result. From the recovered deformation, one can directly compute a number of clinically useful parameters, including strains [29]. We tested the method on four cases (two normals, two patients). A visual inspection of the recovered LV wall and a comparison to manual segmentation suggest that the method is able to successfully recover the deformation of the LV and RV myocardium walls.

## References


1. Tustison, N., Amini, A.: Biventricular Myocardial Strains via Nonrigid Registration of AnFigatomical NURBS Models. IEEE Trans. on Medical Imaging 25 (2006) 94-112
2. Pan, L., Prince, J., Lima, J., Osman, N.: Fast tracking of cardiac motion using 3D-HARP. IEEE Trans. on Biomedical Engineering 52 (2005) 1425-1435
3. Cao, Y., Miller, M., Winslow, R., Younes, L.: Large Deformation Diffeomorphic Metric Mapping of Vector Fields. IEEE Trans. on Medical Imaging 24 (2005) 1216-1230
4. Gilson, W., Yuang, Z., French, B., Epstein, F.: Measurement of myocardial mechanics in mice before and after infarction using multislice displacement-encoded MRI with 3D motion encoding. American Journal of Physiol.- Heart Circ. Physiol. 288 (2005) 1491-1497
5. Meyer, F., Constable, R., Sinusas, A., Duncan, J.: Tracking Myocardial Deformation Using Phase Contrast MR Velocity Fields: A Stochastic Approach. IEEE Trans. on Medical Imaging 15 (1996) 453-465.
6. Kaus, M., Von Berg, J., Weese, J., Niessen, W., Pekar, V.: Automated segmentation of the left ventricle in cardiac MRI. Medical Image Analysis 8 (2004) 245-254
7. Huang, H., Shen, L., Zhang, R., Makedon, F., Hettleman, B., Pearlman, J.:Surface Alignment of 3D Spherical Harmonics Models: Application to Cardiac MRI Analysis. In Medical Image Computing and Computer-Assisted Intervention (2005) 67-74

8. Uzmc, M., Frangi, A., Sonka, M., Reiber, J., Lelieveldt, B.: ICA vs. PCA Active Appearance Models: Application to Cardiac MR Segmentation. In Medical Image Computing and Computer-Assisted Intervention (2003) 451-458
9. Gering, D.: Automatic Segmentation of Cardiac MRI. In Medical Image Computing and Computer-Assisted Intervention (2003) 524-532
10. Van Assen, H., Danilouchkine, M., Behloul, F., Lamb, H., Van Der Geest, R., Reiber, J., Lelieveldt, B.: Cardiac LV Segmentation Using a 3D Active Shape Model Driven by Fuzzy Inference. In Medical Image Computing and Computer-Assisted Intervention (2003) 533-540
11. Montagnat,J., Delingette, H.: 4D deformable Models with temporal constraints: applications to 4D cardiac image segmentation. Medical Image Analysis 9 (2005) 87-100
12. Sermesant, M., Forest, C., Pennec, X., Delingette, H., Ayache, N.: Deformable biomechanical models: Application to 4D cardiac image analysis. Medical Image Analysis 7 (2003) 475-488
13. Remme, E., Augenstein, K., Young, A., Hunter, P.: Parameters Distribution Models for Estimation of Population Based Left Ventricular Deformation Using Sparse Fiducial Markers. IEEE Trans. on Medical Imaging 24 (2005) 381-392
14. Papademteris, X., Sinusas, A., Dione, D., Constable, R., Duncan, J: Estimation of 3-D Left Ventricular Deformation From Medical Images Using Biomedchanical Models. IEEE Trans. on Medical Imaging 21 (2002) 524-532
15. Lorenzo-Valdes, M., Sanchez-Ortiz, G., Mohiaddin, R., Rueckert, D.: Atlas-based segmentation and tracking of 3D cardiac MR images using non-rigid registration. In Medical Image Computing and Computer-Assisted Intervention (2002) 642-650
16. Lorenzo-Valdes, M., Sanchez-Ortiz, G., Elkington, A., Mohiaddin, R., Rueckert, D.: Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. Medical Image Analysis 8 (2004) 255-265
17. Shen D., Sundar, H., Xue, Z., Fan, Y., Litt, H.: Consistent Estimation of Cardiac Motions by 4D Image Registration.In Medical Image Computing and Computer-Assisted Intervention (2005) 67-74
18. Lin, N., Duncan, J.: Generalized Robust Point Matching Using an Exented Free-Form Deformation Model: Application to Cardiac Images. In International Symposium on Biomedical Imaging (2004) 320-324
19. Papademetris, X., Shi. P., Dione, D., Sinusas, A., Duncan, J.: Recovery of soft tissue object deformation using biomechanical models. In Information Processing in Medical Imaging (1999) 352-357
20. Fan L., Chen, C.: LV Motion Estimation Based on the Integration of Continuum Mechanics and Estimation Theory. SPIE Medical Imaging (1999) 81-92
21. Yin F., Chan, C., Judd, R.: Compressibility of Perfused Passive Myocardium. American Journal of Physiol.- Heart Circ. Physiol. 8 (1996) 1864-1870
22. Judd, R., Levy, B.: Effects of Barium-induced Cardiac Contraction on Large and Small Vessel Intramyocardial Blood Volume. Circulation (1991) 217-225
23. Liu, Y., Bahn, R., Ritman, E: Dynamic Intramyocardial Blood Volume: Evaluation with a Radiological Opaque Marker Method. American Journal of Physiol.- Heart Circ. Physiol. 12 (1992) 963-967
24. Vergroesen, I., Noble, M., Spaan, J.:Intramyocardial Blood Volume Change in First Moments of Cardiac Arrest in Anesthetized Goats. American Journal of Physiol.- Heart Circ. Physiol. 4 (1987) 307-316
25. Jakob, M., Hess, O., Jenni, R., Heywood, J., Grimm, J.: Determination of the Left Ventricular Systolic Wall Thickness by Digital Substraction Angiography. European Heart Journal 12 (1991) 573-582

26. Wahaba, G.: Spline Interpolation and Smoothing on the sphere. SIAM Journal Sci. Stat. Comput. 2 (1981) 1-15
27. Carmo, D.: Differential Geometry of Curves and Surfaces. Prentice Hall (1976)
28. Press, W., Flannery, B., Teukolsky, S., Vetterling, W.: Numerical Recipes in C: The Art of Scientific Computing. 2nd ed.Cambrige Univeristy Press (1992)
29. Skrinjar, O., Bistoquet, A.: Cardiac deformation recovery via incompressible transformation decomposition. SPIE Medical Imaging 5747 (2005) 361-370

# A Log-Euclidean Polyaffine Framework for Locally Rigid or Affine Registration

Vincent Arsigny[1], Olivier Commowick[1,2],
Xavier Pennec[1], and Nicholas Ayache[1]

[1] INRIA Sophia - Epidaure Project, 2004 Route des Lucioles BP 93
06902 Sophia Antipolis Cedex, France
Vincent.Arsigny@Polytechnique.org
[2] DOSISoft S.A., 45 Avenue Carnot, 94 230 Cachan, France

**Abstract.** In this article, we focus on the parameterization of non-rigid geometrical deformations with a small number of flexible degrees of freedom . In previous work, we proposed a general framework called *polyaffine* to parameterize deformations with a small number of rigid or affine components, while guaranteeing the invertibility of global deformations. However, this framework lacks some important properties: the inverse of a polyaffine transformation is not polyaffine in general, and the polyaffine fusion of affine components is not invariant with respect to a change of coordinate system. We present here a novel general framework, called *Log-Euclidean polyaffine*, which overcomes these defects. We also detail a simple algorithm, the *Fast Polyaffine Transform*, which allows to compute very efficiently Log-Euclidean polyaffine transformations and their inverses on a regular grid. The results presented here on real 3D locally affine registration suggest that our novel framework provides a general and efficient way of fusing local rigid or affine deformations into a global invertible transformation without introducing artifacts, independently of the way local deformations are first estimated.

## 1 Introduction

The registration of medical images is in general a difficult problem, and numerous methods and tools have been already devised to address this task [9]. Still currently, much effort continues to be devoted to finding adequate measures of similarity, relevant parameterizations of geometrical deformations, efficient optimization methods, or realistic mechanical models of deformations, depending on the precise type of registration considered.

In this article, we focus on the parameterization of non-rigid geometrical deformations with a small number of flexible degrees of freedom . This type of parameterization is particularly well-adapted for example to the registration of articulated structures [11] and to the registration of histological slices [12, 3]. After a *global* affine (or rigid) alignment, this sort of parameterization also allows a finer *local* registration with *very smooth* transformations [5, 10, 6, 13].

In [3], we parameterized deformations with a small number of *rigid or affine components*, which can model smoothly a large variety of local deformations.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 120–127, 2006.

We provided a general framework to fuse these components into a global transformation, called *polyrigid* or *polyaffine*, whose *invertibility* is guaranteed. However, this framework lacks some important properties: the inverse of a polyaffine transformation is not polyaffine in general, and the polyaffine fusion of affine components is not invariant with respect to a change of coordinate system (i.e. is not *affine-invariant*). Here, we present a novel general framework to fuse rigid or affine components, called *Log-Euclidean polyaffine*, which overcomes these defects and yields transformations which can be very efficiently computed.

The sequel of this article is organized as follows. In Section 2, we present the Log-Euclidean polyaffine framework and its intuitive properties. Then, we present the *Fast Polyaffine Transform* (FPT), which allows to compute very efficiently Log-Euclidean polyaffine transformations (LEPTs) and their inverses on a regular grid. Finally, we apply the FPT to a real 3D example, where affine components are estimated with the algorithm of [5]. *Without introducing artifacts*, our novel fusion ensures the invertibility of the global transformation.

## 2 A Log-Euclidean Polyaffine Framework

Before presenting our novel polyaffine framework, let us briefly recall our framework for locally affine registration and the original polyaffine framework, described in [3].

**Locally Affine or Rigid Transformations.** Following the seminal work of [8], we parameterize locally affine (or rigid) transformations by $N$ affine (or rigid) *components*. Each component $i$ consists of an affine transformation $T_i = (M_i, t_i)$ ($M_i$ and $t_i$ are the linear and translation parts) and of a non-negative *weight function* $w_i(x)$, such that the influence of the $i^{\text{th}}$ component at point $x$ is proportional to $w_i(x)$. Here, we assume that the weights are *normalized*: that for all $x$, $\sum_{i=1}^{N} w_i(x) = 1$.

**Direct Fusion of Components.** To obtain a global transformation from several components, the classical approach [14], called here *direct fusion*, simply consists in averaging the associated displacements according to the weights:

$$T(x) = \sum_{i=1}^{N} w_i(x) T_i(x). \tag{1}$$

The transformation obtained using (1) is smooth, but although each component is invertible, the resulting global transformation is *not invertible* in general.

**Previous Polyaffine Framework.** We proposed in [3] to average displacements *infinitesimally*. The resulting global transformation is obtained by integrating an Ordinary Differential Equation (ODE), called polyaffine, which is computationally more expensive, but guarantees the invertibility of global deformations. To define a polyaffine ODE, this approach relies on *principal logarithms* of the *linear parts* $M_i$ of the transformations $T_i$. However, as mentioned in the introduction, this framework lacks some important and desirable properties.

**Logarithm of an Rigid or Affine Transformation.** The key idea of our novel approach is to use the logarithms of the transformations *themselves*. In 3D, the logarithm of an affine (or rigid) transformation $T$ is given in homogeneous coordinates by a 4x4 matrix which is simply the matrix logarithm of the 4x4 matrix representing $T$ [1]:

$$\log(T) = \log \begin{pmatrix} M & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} L & v \\ 0 & 0 \end{pmatrix}.$$

This logarithm is well-defined if and only if none of the eigenvalues of $M$ are non-positive real numbers. See [4] for more details and an efficient numerical algorithm to compute matrix logarithms. Intuitively, this constraint only excludes affine transformations *very far from the identity*, which we did not observe at all in our registration experiments. In particular, for rigid components, this only imposes that (local) rotations be stricly below $\pi$ radians. For a discussion of this limitation, see [1]. In the following, we assume that the logarithms of our affine transformations are well-defined.

**Log-Euclidean Polyaffine Transformations.** Let $(T_i)$ be $N$ affine (or rigid) transformations, and let $(\log(T_i))$ be their logarithms. Using these logarithms, one can fuse the $T_i$ infinitesimally according to the weights $w_i(x)$ with a *stationary* (or *autonomous*) ODE. In homogeneous coordinates, this ODE is the following:

$$\dot{x} = V(x) \stackrel{def}{=} \sum_i w_i(x) \log(T_i).x. \tag{2}$$

The solutions of (2) are always well-defined for all time. The proof is extremely similar to that given in [3]. The value at a point $x$ of the Log-Euclidean polyaffine transformation (LEPT) defined by (2) is given by integrating (2) between time 0 and 1 with $x$ as initial condition. This novel framework is called *Log-Euclidean*, because when the weights $w_i(x)$ do not depend on $x$, the resulting LEPT is simply the affine (or rigid) transformation equal to $\exp(\sum_i w_i \log(T_i))$, i.e. the Log-Euclidean mean of the components, similarly as in our work on tensors [2].

**Remarkable Properties.** The stationarity of (2) yields particularly nice and intuitive properties, conveniently expressed in terms of *flow*. At an instant $s$, the flow $T(s,.)$ of (2) is the mapping which gives the way the ambient space is deformed by the integration of (2) during $s$ units of time. It is always invertible and smooth (as well as its inverse) [16], i.e. it is a *diffeomorphism*.

A classical property of the flow is the following: it is a *one-parameter subgroup* of diffeomorphisms, i.e. $T(s,.) \circ T(t,.) = T(s+t,.)$. Here, it is also a one-parameter subgroup of *LEPTs*. This means that $T(s,.)$ is the $s^{\text{th}}$ power of the Log-Euclidean polyaffine transformation defined by $T(1,.)$. In particular, we have the intuitive properties that the *inverse* of $T(1,.)$ (resp. its square root) is simply $T(-1,.)$ (resp. $T(1/2,.)$), i.e. the LEPT with *identical* weights but whose affine transformations are the *inverses* (resp. square roots) of the original ones. Last but not least, (2) is *affine-invariant*: the Log-Euclidean polyaffine fusion does not depend on the current coordinate system. For more details, see [1].

## 3 Fast Polyaffine Transform

The remarkable (and novel) properties of the Log-Euclidean polyaffine framework allow fast computations of LEPTs. We propose here a very efficient algorithm to evaluate a Log-Euclidean polyaffine transformation on a regular grid.

**Method Overview.** Surprisingly, our fast algorithm generalizes a method widely used to compute *matrix exponentials* to the non-linear case. The basic idea of this method, called 'Scaling and Squaring', is that for a square matrix $M$, we have: $\exp(M) = \exp(\frac{M}{2^N})^{2^N}$. Since the matrix exponential is much simpler to compute for matrices *close to zero*, for example using Padé approximants, one can compute very accurately $\exp\left(\frac{M}{2^N}\right)$ and obtain $\exp(M)$ by squaring recursively $N$ times the result [7]. In the non-linear case, since the flow $T(s,.)$ of (2) is a one-parameter subgroup, we also have:

$$T(1,.) = T\left(\frac{1}{2^N},.\right)^{2^N}, \tag{3}$$

which means that what the deformation observed at time 1 (i.e., the LEPT) results of $2^N$ times the composition of the small deformations observed at time $\frac{1}{2^N}$. Therefore, one can generalize the 'Scaling and Squaring' method to LEPTs in a straightforward way. This method, called the 'Fast Polyaffine Transform' (FPT), follows the same three steps as in the matrix case:

1. **Scaling step:** divide $V(x)$ (speed vectors of (2)) by a factor $2^N$, so that $V(x)/2^N$ is close enough to zero (according to the level of accuracy desired).
2. **Exponentiation step:** $T\left(\frac{1}{2^N},.\right)$ is computed with a numerical scheme.
3. **Squaring step:** using (3), $N$ recursive squarings of $T\left(\frac{1}{2^N},.\right)$ yield an accurate estimation of $T(1,.)$ (only $N$ *compositions of mappings* are used).

**Numerical Scheme for the Exponentiation Step.** Integrating an ODE during a very short interval of time (short with respect to the smoothness of the solution) is quite simple. Generalizing ideas of [3], we use in this article a second-order scheme, called the *affine exponentiation scheme* (A.S.), which is *exact* in the case of a single affine component. It writes in homogeneous coordinates:

$$T\left(\frac{1}{2^N},x\right)_{\text{A.S.}} \stackrel{def}{=} \sum_{i=1}^{N} w_i(x).\exp\left(\frac{1}{2^N}\log\left(T_i\right)\right).x.$$

This choice of scheme comes from our numerical experiments [1] which show that this numerical scheme is on average approximately 40% more accurate than the first-order explicit scheme, with a similar simplicity and computational cost.

**Computational Cost.** An integration of (2) between times 0 and 1 with a time-step of $2^{-N}$ is performed in only $N$ steps, and not in $2^N$ steps as with

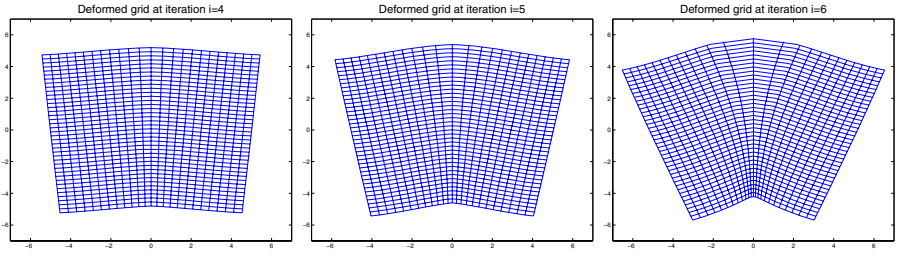



**Fig. 1. Fast polyaffine transform for two rotations: three last squaring steps.** 6 squarings are used in this experiment. Note how the deformation is initially very small, and increases exponentially. The relative accuracy of the estimation of the polyaffine transformation is on average of 0.21%, and the maximal error is below 3.2%.

methods using fixed time-steps. This is somehow comparable with the computational gain obtained by using the Fast Fourier Transform. Interestingly, after the exponentiation step, only $N$ compositions between transformations are needed, which is an operation based on interpolation techniques. In this work, we use bi- and tri-linear interpolations, which are simple and guarantee a continuous interpolation of our transformation.

**Synthetic Experiments.** We measure the accuracy of our results by computing the relative difference with respect to an accurate estimation of the continuous transformation, obtained by a classical integration (i.e., with fixed time step, here $2^{-8}$ ) of (2) for each voxel of the grid, which has $50 \times 40$ pixels. Numerical errors at the boundary of the regular grid are drastically reduced here by adding extra points to the grid so that it contains the boundary of the original grid deformed by direct fusion.

Fig. 1 displays the last 3 squaring steps of a typical FPT, using two rotations of opposite angles of 0.63 radians, (normalized) Gaussian weights ($\sigma = 5$) and a scaling of $2^6$ (i.e., 6 squarings). Errors are low: the relative accuracy of the resulting estimation of the polyaffine transformation is on average 0.21% (instead of approximately 0.6% without an enlarged grid), and the maximal relative error is below 3.2% (instead of 11% wihtout an enlarged grid).

**Inversion with the FPT.** The inverse of a LEPT is simply (and intuitively) the LEPT with the same weights and with inverted affine transformations. Therefore, it can also be computed using the FPT. The accuracy of the inversion is evaluated via the composition of the estimation of the original LEPT and of its inverse by FPT, which should be close to the identity. Fig. 2 shows the evolution of this accuracy when the number of squarings varies, in our example of fusion between two rotations. We thus see that an excellent quality of inversion can be achieved using a small number of squarings, typically 7. The maximal relative error converges below 2% and the mean relative error is of the order of 0.2%. Similar results were obtained in [1] on other examples.

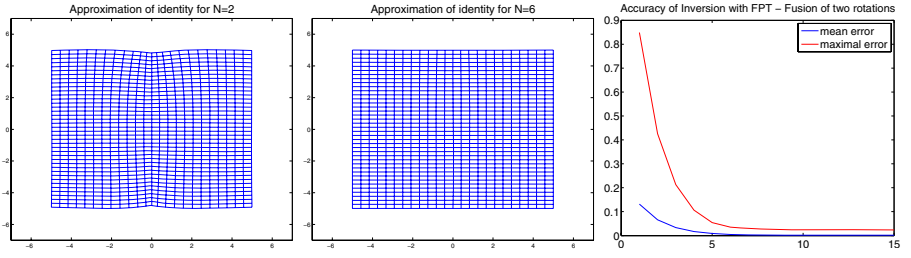



**Fig. 2. Inversion with the FPT. From left to right:** the regular grid deformed by the composition between the FPT of the LEPT and the FPT of its inverse, first with 2 squarings then 6 squarings. **On the right:** evolution of the relative accuracy when the number of squarings varies. An excellent accuracy is already achieved with 6 squarings, and the mean and maximal relative errors converge toward 0.2%and 2%.

## 4 Application to Locally Affine 3D Registration

Let us now consider a real 3D example of locally affine registration, between an atlas of $216 \times 180 \times 180$ voxels and a $T_1$-weighted MR image, with the multi-resolution and robust block-matching algorithm described in [5], without regularization. 7 structures of interest are considered: eyes (1 affine component each), cerebellum (2 components), brain stem (2 components), optic chiasm (1 component), 1 supplementary component (set to the identity) elsewhere. Weight functions are defined in the atlas geometry using mathematical morphology and a smoothing kernel in a preliminary step [5].

**LEPTs as a Post-Processing Tool.** To obtain short computation times (typically 10 minutes), our locally affine registration algorithm estimates affine components using the *direct fusion*. The FPT is used in a *final step* to ensure

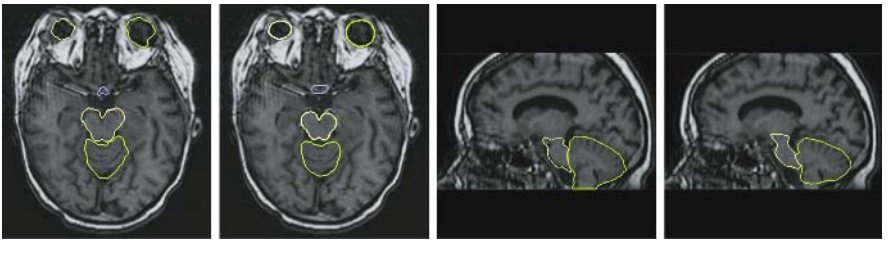


**Fig. 3. Locally affine vs. dense-transformation: smoothness of deformations.** The contours of our structures of interest (eyes, brain stem, cerebellum, optic chiasm) are displayed on the subject and are obtained by deforming those of the atlas using the dense transformation of [15] and using our locally affine framework. **From left to right**: axial slice, with first dense and then locally affine deformations; sagittal slice, with again dense and then locally affine deformations. Note how smoother contours are in the locally affine case, although both accuracies are comparable.

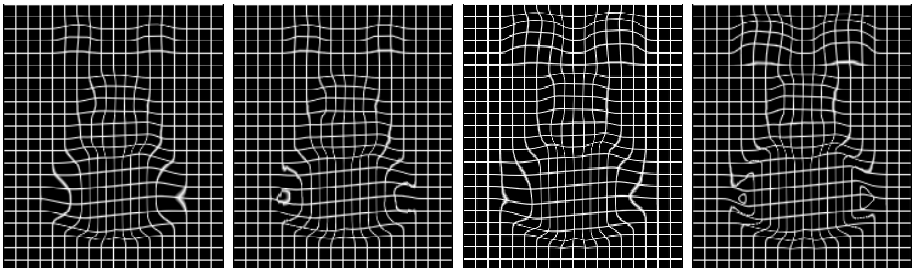


**Fig. 4. Singularity removal with LEPTs.** A 3D regular grid is deformed with the locally affine transformation obtained with the algorithm of [5], two slices are displayed. **From left to right:** polyaffine fusion and direct fusion (for the first slice) and then again polyaffine fusion and direct fusion (second slice). Note how the singularities of the direct fusion disappear with LEPTs. Remarkably, this is obtained without introducing any artifacts: outside singularities, both fusions yield very close results.

the invertibility of the final transformation, as well as to compute its inverse. A typical result of this registration procedure is illustrated by Fig. 3, which shows that the locally affine registration, with much smoother deformations, has an accuracy in the structures of interest which is comparable to the dense transformation case of [15].

Here, the scaling used in $2^8$ and the FPT is computed in $40s$ on a Pentium4 Xeon™2.8 GHz on a $216 \times 180 \times 180$ regular grid. As shown by Fig. 4, the direct fusion of components estimated by [5] can lead to singularities, which is not the case when the FPT is used. Remarkably, both fusions are very close *outside* of regions with singularities. This means that no artifacts are introduced by the FPT, which justifies *a posteriori* the estimation of affine components with the (faster) direct fusion.

## 5 Conclusion and Perspectives

In this work, we have presented a novel framework to fuse rigid or affine components into a global transformation, called *Log-Euclidean polyaffine*. Similarly to the previous polyaffine framework of [3], it guarantees the *invertibility* of the result. However, contrary to the previous framework, this is achieved with very intuitive properties: for example the inverse of a LEPT is a LEPT with identical weights and inverted affine components. Moreover, this novel fusion is *affine-invariant*, i.e. does not depend on the choice of coordinate system. We have also shown that remarkably, and contrary to previous polyaffine transformations, the specific properties of LEPTs allow their fast computations on regular grids, with an algorithm called the 'Fast Polyaffine Transform', whose efficiency is somehow comparable to that of the Fast Fourier Transform.

In the example of locally affine 3D registration presented here, we use LEPTs in a final step to fuse the affine components estimated during the algorithm of

[5]. With the FPT, this is done very efficiently. Remarkably, the novel fusion is *very close* to the direct fusion in regions without singularities. This suggests that our novel framework provides a general and efficient way of fusing local rigid or affine deformations into a global invertible transformation without introducing artifacts, *independently* of the way local affine deformations are first estimated.

## References


1. V. Arsigny, O. Commowick, X. Pennec, and N. Ayache. A fast and Log-Euclidean polyaffine framework for locally affine registration. Research report RR-5865, INRIA, March 2006.
2. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Fast and simple calculus on tensors in the Log-Euclidean framework. In *MICCAI (1)*, pages 115–122, 2005.
3. V. Arsigny, X. Pennec, and N. Ayache. Polyrigid and polyaffine transformations: a novel geometrical tool to deal with non-rigid deformations - application to the registration of histological slices. *Med. Im. Anal.*, 9(6):507–523, December 2005.
4. S. Hun Cheng, N. J. Higham, C. S. Kenney, and A. J. Laub. Approximating the logarithm of a matrix to specified accuracy. *SIAM J. Matrix Anal. Appl.*, 22(4):1112–1125, 2001.
5. O. Commowick, V. Arsigny, J. Costa, G. Malandain, and N. Ayache. An efficient multi-affine framework for the registration of anatomical structures. In *Proceedings of ISBI'2006*. IEEE, 2006. To appear.
6. A. Cuzol, P. Hellier, and E. Mémin. A novel parametric method for non-rigid image registration. In *Proc. of IPMI'05*, number 3565 in LNCS, pages 456–467, 2005.
7. N. J. Higham. The scaling and squaring method for the matrix exponential revisited. *SIAM J. Matrix Anal. Appl.*, 26(4):1179–1193, 2005.
8. J.A. Little, D.L.G. Hill, and D.J. Hawkes. Deformations incorpotationg rigid structures. *CVIU*, 66(2):223–232, May 1996.
9. J.B.A. Maintz and M.A Viergever. A survey of medical registration. *Medical image analysis*, 2(1):1–36, 1998.
10. R. Narayanan, J.A. Fessler, H. Park, and C.R. Meyer. Diffeomorphic nonlinear transformations: A local parametric approach for image registration. In *Proceedings of IPMI'05*, volume 3565 of LNCS, pages 174–185, 2005.
11. X. Papademetris, D.P. Dione, L.W. Dobrucki, L.H. Staib, and A.J. Sinusas. Articulated rigid registration for serial lower-limb mouse imaging. In *MICCAI'05 (2)*, pages 919–926, 2005.
12. A. Pitiot, E. Bardinet, P.M. Thompson, and G. Malandain. Piecewise affine registration of biological images for volume reconstruction. *Med. Im. Anal.*, 2005. accepted for publication.
13. 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 Trans. Medecal Imaging*, 18(8):712–721, 1999.
14. D. Sheppard. A two-dimensionnal interpolation function for irregularly spaced data. In *23rd National Conference of the ACM*, pages 517–524, 1968.
15. R. Stefanescu, X. Pennec, and N. Ayache. Grid powered nonlinear image registration with locally adaptive regularization. *Med. Im. Anal.*, 8(3):325–342, September 2004.
16. M. Tenenbaum and H. Pollard. *Ordinary Differential Equations*. Dover, 1985.

# Introduction to the Non-rigid Image Registration Evaluation Project (NIREP)

Gary E. Christensen[1], Xiujuan Geng[1], Jon G. Kuhl[1], Joel Bruss[2], Thomas J. Grabowski[2,3], Imran A. Pirwani[4], Michael W. Vannier[5], John S. Allen[6,7], and Hanna Damasio[6]

[1]Department of Electrical and Computer Engineering, The University of Iowa
[2]Department of Neurology, The University of Iowa
[3]Department of Radiology, The University of Iowa
[4]Department of Computer Science, The University of Iowa
[5]Department of Radiology, University of Chicago
[6]The Dornsife Cognitive Neuroscience Imaging Center, Univ. of Southern California
[7]Department of Anthropology, Univ. of Southern California

**Abstract.** Non-rigid image registration (NIR) is an essential tool for morphologic comparisons in the presence of intra- and inter-individual anatomic variations. Many NIR methods have been developed, but are especially difficult to evaluate since point-wise inter-image correspondence is usually unknown, i.e., there is no "Gold Standard" to evaluate performance. The Non-rigid Image Registration Evaluation Project (NIREP) has been started to develop, establish, maintain, and endorse a standardized set of relevant benchmarks and metrics for performance evaluation of nonrigid image registration algorithms. This paper describes the basic framework of the project.

## 1 Introduction

Image registration is important for many applications, including longitudinal evaluations in individuals, comparison between individuals, creation of population atlases, use of atlas-linked information in individual cases, delivery of precision therapies, and many others. Non-rigid image registration is a more general approach than the widely used affine and rigid methods, but requires more complex methodology and computational effort to implement. Evaluating the performance of non-rigid image registration algorithms is a difficult task since point-wise correspondence from one image to another is not unique. That is there is rarely if ever a ground truth correspondence map to judge the performance of a registration algorithm.

We have started the Non-rigid Image Registration Evaluation Project (NIREP) to develop software tools and provide shared image validation databases for rigorous testing of non-rigid image registration algorithms. NIREP will extend the scope of prior validation projects by developing evaluation criteria and metrics using large image populations, using richly annotated image databases, using computer simulated data, and increasing the number and types of evaluation criteria.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 128–135, 2006.

The goal of this project is to establish, maintain, and endorse a standardized set of relevant benchmarks and metrics for performance evaluation of nonrigid image registration algorithms. Furthermore, these standards will be incorporated into an exportable computer program to automatically evaluate the registration accuracy of nonrigid image registration algorithms.

## 2 Prior Studies

To date, few attempts have been made to objectively evaluate and compare the performance of image registration algorithms using standard evaluation criteria. Two projects that stand out in this regard are the "Retrospective Image Registration and Evaluation Project" [1] led by J. Michael Fitzpatrick of Vanderbilt University for evaluating multimodality rigid registration accuracy and the non-rigid registration evaluation project entitled "Retrospective Evaluation of Inter-subject Brain Registration" [2] led by Christian Barillot of IRISA/INRIA-CNRS Rennes, France. In both of these projects, a common set of images was used to evaluate the performance of registration algorithms. Developers from around the world participated in these projects by registering the images with their own registration algorithms and sending the resulting transformations back to the home site for analysis. The benefits of involving external participants include eliminating implementation biases, distributing the processing load, and providing an incentive to produce good results.

Another important validation/evaluation project is the VALMET software tool for assessing and improving 2D and 3D object segmentation developed by Guido Gerig et al. [3] (www.ia.unc.edu/public/valmet/). The VALMET software was the first publicly available software tool for measuring and visualizing the differences between multiple corresponding medical image segmentations. It includes four algorithms for comparing segmentations: overlap ratio, Haussdorf distance, surface distance, and probabilistic overlap. The NIREP software evaluates image registration algorithm performance similar to the way the VALMET software evaluates image segmentation performance. In the future, all the VALMET evaluation metrics will be incorporated in to the NIREP software since automatic image segmentation produced from image registration is often used to evaluate registration performance.

## 3 NIREP Framework

The Nonrigid Image Registration Evaluation Project (NIREP) will build upon these earlier projects and provide a rich set of metrics for evaluating registration performance. We are building an exportable software tool that has nonrigid image registration evaluation metrics built into it. This is in contrast to previous projects that collected registration transformations from external participants and performed the analysis internally.

Under our model, NIREP software users will process their own data and evaluate the performance of different nonrigid registration algorithms on this

data. The tests will be fixed in the program such that the users will not be able to affect the results. This model has the advantage of standardizing the evaluation metrics and distributing the processing load. It will provide researchers with a tool to compare the performance of multiple registration algorithms on their own data so they can make an informed decision regarding the best algorithm for their specific application. It also provides researchers with a tool to validate their research results. The full disclosure of all validation tests performed as a result of NIREP project will be collected and analyzed. The results and analysis will be disseminated through publications and a central web site (www.nirep.org).

Evaluating the performance of nonrigid image registration algorithms is a difficult task since point-wise correspondence from one image to another is not known. In the absence of the truth, a diverse set of evaluation metrics will be used to evaluate registration performance on multiple well documented image databases. These tests evaluate the performance of image registration algorithms with respect to their transformation properties, agreement with human experts, agreement with computer simulated data, and other indirect performance tests. Complete and accurate reporting of validation tests will be published so that others will be able to detect potential bias (internal validity) and assess the generalizability and applicability of the results (external validity).

The NIREP software evaluates image registration algorithms in a similar fashion to the model used to evaluate the performance of computer hardware and systems. Computer systems are evaluated and compared using not one but many diverse criteria. The diversity of the tests serves both to evaluate system performance and to demonstrate the trade-offs between various systems. Popular examples of groups that evaluate computer hardware performance include the Standard Performance Evaluation Corporation (www.specbench.org) and Tom's Hardware (www.tomshardware.com). The diversity of the image registration metrics developed in NIREP evaluate registration algorithm performance and demonstrate trade-offs between different algorithms.

## 4 Evaluation Database

Our initial evaluation database consists of a population of 16 richly annotated 3D MR image volumes corresponding to 8 normal adult males and 8 females. These data sets were selected from a database of healthy right-handed individuals acquired in the Human Neuroanatomy and Neuroimaging Laboratory, The University of Iowa, and now part of the Laboratory of Computational Neuroimaging, The University of Iowa. The males have a mean age of 32.5 years, standard deviation of 8.4 years and range in age from 25 to 48. The females have a mean age of 29.8 years, standard deviation of 5.8 and range in age from 24 to 41. The complete population will be used to evaluate the non-rigid image registration performance for complexly shaped neuroanatomical structures.

The 16 MR data sets have been segmented into 32 gray matter regions of interest (ROIs). Fig. 1 shows an example of the segmentations associated with the MR data sets. Some of these data sets and their segmentations appear in a

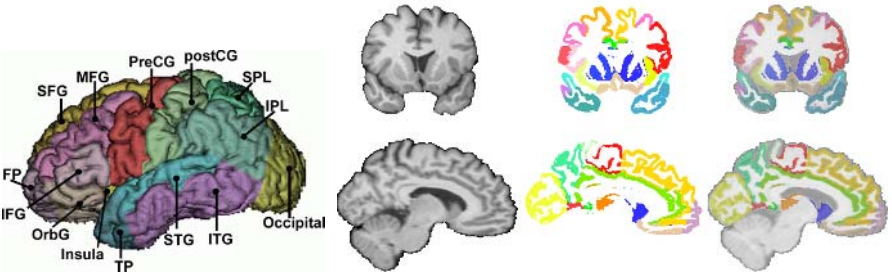



**Fig. 1.** Shown are the segmentations available in the neuroanatomy 1 evaluation database. The segmentations include gray matter regions in the frontal, parietal, temporal, and occipital lobes; cingulate gyrus; and insula. The cerebellum, hypothalamus, and brain stem are currently not segmented currently.

new atlas by Hanna Damasio that illustrates the normal neuroanatomy of the human brain, and have been used in several publications about the morphometric analysis of the normal human brain [4, 5, 6].

The brains were initially segmented with Brainvox [7] using the criteria described in papers by John Allen et al. [4, 5, 6]. The resulting segmentations partitioned the brain in to regions that contained both gray and white matter. Although the image volumes were carefully segmented, the segmentation process was done in 2D. As a result, the segmentations were smooth in the plane of segmentation but had rough edges when viewed in oblique slices. In addition, many of the segmentations had to have arbitrary boundaries within the white matter since region boundaries are well defined at sulci level but have to rely on "connecting lines" between the depth of the sulci within the white matter. These initial segmentations were then restricted to the gray matter to fix the boundary problems in the white matter. Gray matter segmentations were generated using the approach described in Grabowski et al. [8]. The gray matter segmentations were applied to the regional segmentations to remove the white matter from the segmentations. This produced gray matter ROIs with smooth boundaries at the outer surface of the cortex and at the gray/white interface. The segmentations were then hand edited using the Analyze$^{TM}$ software (Mayo Clinic, Rochester Minnesota) to produce the final gray matter segmentations.

## 5 Evaluation Criteria

Since there is rarely if ever a "Gold Standard" to evaluate image registration results, no metric alone is sufficient to evaluate the performance of a nonrigid registration algorithm. However, using information from many different diverse metrics will provide a good indication of the non-rigid image registration performance. In this initial phase of the project, we have included the following four metrics: (1) relative overlap, (2) intensity variance, (3) inverse consistency and (4) transitivity.

***Relative Overlap Metric:*** The alignment of the regions of interest (ROIs) is a good indicator of how well two images are registered. The relative overlap of the segmentations is a measure of how well two corresponding segmented regions agree with each other. The Relative Overlap metric is given by $RO(P,S) = \frac{volume(P \cap S)}{volume(P \cup S)}$ where $P$ and $S$ are two corresponding segmentations. In the context of image registration, $P$ corresponds to a segmentation transformed from image $i$ to $j$ compared to the corresponding segmentation $S$ defined in image $j$.

The NIREP software generates a variety of images for evaluation of local spatial registration performance and summarizes relevant performance statistics for each region of interest with tables and plots. Figure 2 shows an example of relative overlap performance for the superior frontal gyrus (SFG) and Table 1 shows statistics for the relative overlap before and after registration for the left and right SFG (LSFG, RSFG), the left and right temporal pole (LTP, RTP), and the left and right orbito-frontal gyrus (LOFG, ROFG). As Fig. 2 illustrates, images provide the user with visual feedback for evaluating how well the superior frontal gyrus was registered before and after registration. Similarly, numerical data such as shown in Table 1 provides the user with quantitative feedback of algorithmic performance. Note that the 32 gray matter segmentations currently included in the neuroanatomy 1 database have relatively small volumes compared to the brain volume and have large surface areas compared to their volume. Segmentations with small volumes and large surface areas typically have small relative overlaps even for good registrations. These segmentations should provide a challenging goal for registration.

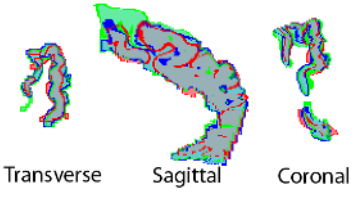



**Fig. 2.** Relative overlap images for the right Superior Frontal Gyrus before and after intersubject registration. Red - target, Green - before registration, Blue - after registration.

***Intensity Variance Metric:*** A common method used to measure image registration performance is to register a population of images with a target image and average the intensities of the registered images. The idea is that the better the registration algorithm is, the closer each registered image looks to the target image and the sharper the intensity average image. One way to measure the sharpness of the intensity average image is to compute the variance of the registered intensity images. The voxel-wise intensity variance (IV) of a population of $M$ images registered to image $j$ is computed as

$$IV_j(x) = \frac{1}{M-1}\sum_{i=1}^{M}(T_i(h_{ij}(x)) - ave_j(x))^2 \quad \text{where} \quad ave_j(x) = \frac{1}{M}\sum_{i=1}^{M} T_i(h_{ij}(x)), \tag{1}$$

$T_i$ is the $i^{th}$ image of the population and $h_{ij}(x)$ is the transformation from image $i$ to $j$ with respect to a Eulerian coordinate system.

**Table 1.** Relative overlap (RO) summary table for 132 ($2\times(11+\cdots+1)$) transformations registered a population of 12 out of the 16 data sets using small deformation inverse-consistent linear-elastic (SICLE) registration method [9]. The voxel size is $0.7 \times 0.7 \times 0.7$ mm$^3$. Due to lack of space, we only report the RO of 6 regions of interest (ROIs) as a representative sample of the RO of all 32 ROIs. The 6 ROIs include: Left/Right Temporal Poles (LTP, RTP), Left/Right Superior Frontal Gyrus (LSFG, RSFG), and Left/Right Orbital Frontal Gyrus (LOFG, ROFG).

| | LTP | RTP | LSFG | RSFG | LOFG | ROFG |
|---|---|---|---|---|---|---|
| Ave Vol ($voxel^3$) | 27900 | 30900 | 79600 | 76700 | 45700 | 47300 |
| Std Dev Vol | 4850 | 4320 | 9960 | 12200 | 3250 | 5080 |
| Ave Surface Area ($voxel^2$) | 11100 | 12200 | 36900 | 37000 | 21800 | 22600 |
| Std Dev Surface Area | 1700 | 1360 | 4160 | 4700 | 1700 | 2610 |
| Ave RO: before registration | 0.365 | 0.237 | 0.267 | 0.263 | 0.338 | 0.194 |
| Ave RO: after SICLE | 0.542 | 0.334 | 0.364 | 0.360 | 0.469 | 0.290 |
| RO Std dev: before registration | 0.0733 | 0.0632 | 0.0399 | 0.0613 | 0.0747 | 0.0505 |
| RO Std dev: after SICLE | 0.0503 | 0.0495 | 0.0341 | 0.0438 | 0.0295 | 0.0460 |

***Inverse Consistency Metric:*** The inverse consistency metrics evaluates registration performance based on desired transformation properties [9, 10, 11]. The inverse consistency metric measures the inverse consistency error between a forward and reverse transformation between two images. Ideally the forward transformation equals the inverse of the reverse transformation implying a consistent definition of correspondence between two images, i.e., correspondence defined by the forward transformation should be the same as that defined by the reverse transformations. Thus, composing the forward and reverse transformations together produces the identity map when there is no inverse consistency error. The inverse consistency error is defined as the squared difference between the composition of the forward and reverse transformations and the identity mapping.

The voxel-wise cumulative inverse consistency error (CICE) with respect to template image $j$ is computed as

$$CICE_j(x) = \frac{1}{M}\sum_{i=1}^{M} ||h_{ji}(h_{ij}(x)) - x||^2 \tag{2}$$

where $h_{ij}$ is the transformation from image $i$ to $j$, $M$ is the number of images in the evaluation population and $||\cdot||$ is the standard Euclidean norm.

The CICE is an example of a necessary evaluation metric for evaluating registration performance but is not a sufficient evaluation metric to guarantee good correspondence. For example, two identity transformations have zero inverse consistency error but in general provide poor correspondence between two images. However, a pair of transformations that provide good correspondence between images should have zero CICE.

***Transitivity Metric:*** The transitivity metric [11] evaluates how well all the pairwise registrations of the image population satisfy the transitivity property.

The transitivity property is important to minimize correspondence errors when two transformations are composed together. Ideally, transformations that define correspondence between three images should project a point from image $A$ to $B$ to $C$ to $A$ back to the original position. The transitivity error for a set of transformations is defined as the squared error difference between the composition of the transformations between three images and the identity map.

The voxel-wise cumulative transitivity error (CTE) with respect to template image $j$ is computed as

$$CTE_k(x) = \frac{1}{(M-1)(M-2)} \sum_{\substack{i=1 \\ i \neq k}}^{M} \sum_{\substack{j=1 \\ j \neq i \\ j \neq k}}^{M} ||h_{ki}(h_{ij}(h_{jk}(x))) - x||^2. \qquad (3)$$

The CTE is another example of a necessary evaluation metric but is not a sufficient evaluation metric for guaranteeing good correspondence. For example, a set of transformations that have zero transitivity error does not imply good correspondence as can be see with a set of identity transformations. However, a set of transformations that provide good correspondence between images in a population should have zero transitivity error.

## 6 Future Work

In the future, additional evaluation databases and metrics will be added to NIREP. The NIREP project will serve as a repository for evaluation databases and metrics developed at The University of Iowa and contributed from researchers at other institutions. Extensive documentation will be collected describing the strengths and weaknesses of evaluation databases, metrics, and registration algorithms. The NIREP project, with the help of the medical imaging community, will establish, maintain, and endorse a standardized set of relevant benchmarks and metrics for performance evaluation of nonrigid image registration algorithms.

## Acknowledgments

This work was supported by NIH grants EB004126, CA096679 and HL64368.

## References

1. Jay West, J. Michael Fitzpatrick, et al., "Comparison and evaluation of retrospective intermodality brain image registration techniques", vol. 21, pp. 554–566, 1997.
2. P. Hellier, C. Barillot, L. Corouge, B. Gibaud, G. Le Goualher, D.L. Collins, A. Evans, G. Malandain, N. Ayache, G.E. Christensen, and H.J. Johnson, "Retrospective evaluation of inter-subject brain registration", *IEEE Transactions on Medical Imaging*, vol. 22, pp. 1120–1130, 2003.

3. Guido Gerig, Matthieu Jomier, and Miranda Chakos, "Valmet: A new validation tool for assessing and improving 3d object segmentation", in Wiro J. Niessen and Max A. Viergever, editors, *MICCAI 2001*, vol. LNCS 2208, pp. 516–528. Springer, 2001.
4. Hanna Damasio, *Human brain anatomy in computerized images*, Oxford University Press, New York, 2nd edition, 2005.
5. Allen JS, Damasio H, and Grabowski TJ, "Normal neuroanatomical variation in the human brain: an mri-volumetric study", *Am J Phys Anthropol*, vol. 118, pp. 341–58, 2002.
6. Allen JS, Damasio H, Grabowski TJ, Bruss J, and Zhang W, "Sexual dimorphism and asymmetries in the gray white composition of the human cerebrum", *NeuroImage*, vol. 18, pp. 880–894, 2003.
7. R.J. Frank, H. Damasio, and T.J. Grabowski, "Brainvox: an interactive, multimodel, visualization and analysis system for neuroanatomical imaging", *NeuroImage*, vol. 5, pp. 13–30, 1997.
8. T.J. Grabowski, R.J. Frank, N.R. Szumski C.K. Brown, and H. Damasio, "Validation of partial tissue segmentation of single-channel magnetic resonance images of the brain", *NeuroImage*, vol. 12, pp. 640–656, 2000.
9. G.E. Christensen and H.J. Johnson, "Consistent image registration", vol. 20, pp. 568–582, July 2001.
10. H.J. Johnson and G.E. Christensen, "Consistent landmark and intensity-based image registration", vol. 21, pp. 450–461, 2002.
11. G.E. Christensen and H.J. Johnson, "Invertibility and transitivity analysis for nonrigid image registration", *Journal of Electronic Imaging*, vol. 12, pp. 106–117, Jan. 2003.

# A Unified Framework for Atlas Based Brain Image Segmentation and Registration

E. D'Agostino, F. Maes*, D. Vandermeulen, and P. Suetens

Katholieke Universiteit Leuven, Faculties of Medicine and Engineering, Medical Image Computing (Radiology - ESAT/PSI), University Hospital Gasthuisberg, Herestraat 49, B-3000 Leuven, Belgium
Emiliano.DAgostino@uz.kuleuven.ac.be

**Abstract.** We propose a unified framework in which atlas-based segmentation and non-rigid registration of the atlas and the study image are iteratively solved within a maximum-likelihood expectation maximization (ML-EM) algorithm. Both segmentation and registration processes minimize the same functional, i.e. the log-likelihood, with respect to classification parameters and the spatial transformation. We demonstrate how both processes can be integrated in a mathematically sound and elegant way and which advantages this implies for both segmentation and registration performance. This method (Extended EM, EEM) is evaluated for atlas-based segmentation of MR brain images on real data and compared to the standard EM segmentation algorithm without embedded registration component initialized with an affine registered atlas or after registering the atlas using a mutual information based non-rigid registration algorithm (II).

## 1 Introduction

Atlas-based image segmentation and atlas-to-image non-rigid registration are two challenging problems that are currently actively being investigated in medical image analysis. Traditionally these two problems have been solved independently from each other, although it is clear that a duality exists between the two processes and that they can mutually help each other. In the atlas-based MR brain image segmentation approach proposed in [11], (affine) image registration is applied to match a digital brain atlas to the study image to be segmented, which is used to initialize the segmentation process. The same atlas is used during segmentation iterations to spatially constrain the classification process. Hence, the transformation used to align the atlas to the study image has important impact on the final segmentation result.

On the other hand image registration can also benefit from available segmentation results. In a traditional mutual information based registration approach, images are registered based on intensity information only. Recently, several approaches have been presented [12, 8, 9] that incorporate voxel label information in the registration criterion.

* Frederik Maes is Postdoctoral Fellow of the Fund for Scientific Research - Flanders (FWO-Vlaanderen, Belgium).

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 136–143, 2006.

At the best of our knowledge, only three authors have addressed the topic of merging voxel based brain segmentation and registration [13, 1, 4]. In [13] the authors formulate the joint segmentation/registration problem as the Maximum A Posteriori (MAP) estimation of the segmentation S and the registration transformation T. The authors define one similarity measure for the registration and one for the segmentation component. The joint problem is then solved by finding the solution to the combined registration and segmentation similarity measures. The entire framework is formulated for rigid registration and no bias correction is performed during the segmentation step. In [4] the authors extend the work of [13] to include a B-spline based non-rigid transformation and an hidden Markov random field to improve the segmentation performance.

A similar method is proposed by Ashburner et al. [1]. However, our method differs by the choice of the deformation model, using a viscous fluid regularization that allows for larger deformations than the discrete cosine representation of Ashburner et al [1]. Moreover, in this paper we present an extensive validation of our combined registration and segmentation approach in the context of atlas-based MR brain image segmentation and demonstrate that the simultaneous solution of both problems within a single mathematical framework significantly improves the segmentation performance over a sequential application of the non-rigid registration and segmentation steps.

## 2 Method

In [11], an approach for automated intensity-based tissue classification of MR brain images is introduced. This method uses the Expectation-Maximization algorithm to iteratively estimate the parameters $\theta = (\mu_k, \theta_k, k = 1...n)$ of a Gaussian mixture model (assuming the intensities of each tissue class to be normally distributed with unknown mean and variance, but corrupted by a spatially varying intensity inhomogeneity or bias field) and simultaneously classify each voxel accordingly, such as to maximize the likelihood $p(I|\theta)$ of the intensity data $I$ given the model. The method (in what follows simply called EM) is initialized by providing initial tissue classification maps for white matter (WM), gray matter (GM), cerebrospinal fluid (CSF) and OTHER derived from a digital brain atlas after appropriate spatial normalization of the atlas to the study images. However, the atlas is not only used to initialize the EM procedure, but also serves as a spatially varying prior that constrains the classification during parameter estimation and in the final classification step. The probabilistic tissue classification $L$ is obtained as the a posteriori probability of tissue label given the observed image intensity and the estimated intensity parameters, which, assuming that all voxels are independent, is computed using Bayes' rule as $p(l_k = j|I_k, \theta) \propto p(I_k|l_k = j, \theta).p(l_k = j)$ with $l_k$ the label assigned to voxel $k$, $j$ the various tissue classes, $I_k$ the intensity of voxel $k$, $p(I_k|l_k = j, \theta)$ the probability of the observed intensity given the specified class label as derived from the Gaussian mixture model and $p(l_k = j)$ the prior probability of voxel $k$ to belong to class $j$, which is simply the atlas registered to the image $I$. Hence, the quality

of the atlas-to-image registration has a direct impact on the segmentation result through the above relation and the impact of the atlas model ($p(l_k = j)$) is as important as that of the intensity data ($p(I_k|l_k = j, \theta)$) itself.

Formally, in the EM approach the log-likelihood function $Q$ is defined as:

$$Q(\theta|\theta^m) = \sum_i \sum_k p_{ik}^{m+1}.(\log p(l_i = k) + \log p(y_i^{'}|l_i = k, \mu_k, \sigma_k, C)) \quad (1)$$

with $p_{ik}^{m+1}$ the a posteriori probability for voxel $i$ to belong to class $k$ (i.e. the tissue classification constructed by segmentation), $p(l_i = k)$ the a priori probability (derived from the brain atlas) and $p(y_i^{'}|l_i = k, \mu_k, \sigma_k, C)$ the Gaussian class-specific intensity distribution with parameters $\mu$ and $\sigma$ for the bias-field corrected intensity $y^{'}$ with bias field parameters C. (We ignore the bias field in the following derivation, although the bias field correction component is included in all experiments in this paper, see [11]). In the Expectation step of the EM algorithm, the parameters of the Gaussian mixture model are estimated iteratively based on the current tissue classification:

$$\frac{\partial Q}{\partial \mu_k^{m+1}} = 0 \Longrightarrow \mu_k^{m+1} = \frac{\sum_{i=1}^n p_{ik}^{(m)}.y_i}{\sum_{i=1}^n p_{ik}^{(m)}} \quad (2)$$

$$\frac{\partial Q}{\partial \sigma_k^{m+1}} = 0 \Longrightarrow (\sigma_k^{m+1})^2 = \frac{\sum_{i=1}^n p_{ik}^{(m)}(y_i - \mu_k^{m+1})^2}{\sum_{i=1}^n p_{ik}^{(m)}}$$

In the Maximization step, voxel classification is performed based on the new estimate of the parameters:

$$p_{ik}^{(m+1)} = \frac{p(y_i|l_i = k, \mu_k, \sigma_k).p(l_i = k)}{\sum_j p(y_i|l_i = j, \mu_j, \sigma_j).p(l_i = j)} \quad (3)$$

The method iterates between both steps until convergence.

In the work presented here, we extend this approach to take into account a spatial transformation mapping the (atlas) prior $p(l_i = k)$ onto the space of the study image. More specifically, starting from the same likelihood function, the E-step of the EM algorithm is extended as follows:

$$\frac{\partial Q}{\partial T} = \sum_i \sum_k \frac{\partial Q}{\partial p(l_k = i)} \frac{\partial p(l_k = i)}{\partial T}$$

where $T$ represents the transformation parameter.

If we use trilinear interpolation to warp the atlas to the study image according to the transformation $T$ (but any other differentiable interpolation scheme can also be used), we have:

$$p(l_i = k|T) = \sum_{j=1}^n w_{ij} p_{j,k} \Longrightarrow \frac{\partial p(l_k = i|T)}{\partial T} = \sum_{j=1}^n \frac{\partial w_{ij}}{\partial T} p_{j,k}$$

where $n$ is the number of neighbors (8 in three dimensions), $p_{j,k}$ is the probability of neighbor $j$ to belong to class $k$ and $w_{ij}$ are the trilinear interpolation weights. As it can be seen the derivative $\frac{\partial p(l_k=i)}{\partial T}$ of $p(l_i = k)$ with respect to $T$ can simply be computed by differentiating the trilinear weights. It follows therefore:

$$\frac{\partial Q}{\partial T} = \sum_i \sum_k \frac{p_{ik}^{m+1}}{p(l_i = k|T)} \sum_{j=1}^{n} \frac{\partial w_{ij}}{\partial T} p_{j,k} \tag{4}$$

We need to solve for the deformation for which $\frac{\partial Q}{\partial T} = 0$. While for $\mu$ and $\sigma$ a closed form solution could be derived, solving for $T$ requires an iterative procedure. We interpretate $\frac{\partial Q}{\partial T}$ as a force field that can be defined in each voxel by considering individual voxel displacements. Akin to our previous registration schemes based on mutual information [7] and label-based similarity measures [8], we use this force field to drive the registration by replacing the force field in our previous schemes by the expression for $\frac{\partial Q}{\partial T}$ above. A viscous fluid regularizer was used.

Hence, the segmentation and registration processes are linked and collaborate through the joint optimization of the same functional $Q$.

## 3 Results

We performed two different experiments. In the first experiment we investigated the behavior of the algorithm, in normal conditions, by segmentating 20 (normal) brain images. These images are part of a database ([2]) including T1 images and probability labels for 49 regions in the brain. The different labels were obtained by manual delineation and have been used as ground truth.

In the second experiment we studied the feasibility of our algorithm in presence of gross morphological distortions as they occur in PeriVentricular Leukomalacia (PVL) child brain images. Here no ground truth was present and therefore a qualitative (visual) analysis of the results was used.

In both experiments we used an home-made brain atlas as floating image ([3]). This atlas is composed by a T1 image, priors for white matter matter, gray matter, csf and max-probability labels for 49 regions in the brain.

### 3.1 Normal Brain Images Database

In this first experiment we applied the EEM algorithm (joint segmentation and registration) to each of the 20 images. Our atlas was used to initialize the segmentation and was iteratively deformed onto the study image. As a result of the EEM algorithm we obtained segmentation maps for WM, GM and CSF and deformation fields. In order to assess the results of the EEM algorithm we deformed the atlas using the II algorithm onto each of the 20 images.

We compared the quality of the obtained registrations by computing overlap coefficients between the ground truth for WM, GM and CSF (manual segmentations in the database) and the corresponding atlas priors after affine registration and after deformation with the EEM and the II algorithms.

We also compared the segmentation issued from the EEM algorithm with those obtained by applying the EM algorithm initialized with the atlas affine registered (EM-A) or non-rigid deformed onto each of the study images (EM-II). In this was way we intended to evaluate the impact of the registration (affine or non-rigid) on the segmentation.

Tables 1 and 2 show mean (over the 20 images) overlap coefficients for WM, GM and CSF for the EEM, EM-A and EM-II methods. Figure 1 shows the results in a bar-plot form.

**Table 1.** Overlap coefficients (mean values) (in %) between the ground truth and the atlas registered to the study image (affine (EM-A), integrated segmentation and registration(EEM) and non rigid (EM-II)

| | EM-A | EEM | EM-II |
|---|---|---|---|
| WM | 87.24 | 91.24 | 91.25 |
| GM | 88.74 | 92.47 | 91.95 |
| CSF | 75.54 | 85.18 | 80.97 |

**Table 2.** Overlap coefficients (mean values) (in %) between the ground truth and the segmentation maps obtained using different algorithms (EM-A, EEM and EM-II)

| | EM-A | EEM | EM-II |
|---|---|---|---|
| WM | 72.55 | 84.55 | 84.19 |
| GM | 77.19 | 86.89 | 86.29 |
| CSF | 49.35 | 69.39 | 63.00 |

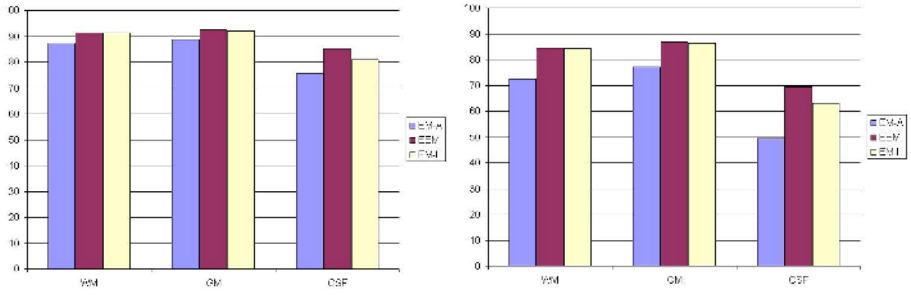



**Fig. 1.** Overlap coefficients for the EM-A the EEM and the EM-II methods between the ground truth and the registered atlas (affine or non-rigid) (left) and between the ground truth and the computed segmentation maps (right)

As it can be observed the use of integrated segmentation and registration leads to better overlaps results, in particular for CSF. In fact not only a better

registration is achieved (with respect to overlap coefficients), as shown in table 2, but also the segmentation maps obtained are improved (table 1).

### 3.2 PVL Brain Image

In this second experiment we had no ground truth. The results could therefore only be evaluated in a qualitative way. Nonetheless simple visual inspection showed the importance of combining registration and segmentation.

In figure 2 we show the segmentation results obtained by using an affine registered atlas (EM-II) and by using the EEM method. Although just in a qualitative way, it's clear that in this case the results generated by the standard EM method are incorrect. As a result the CC registration algorithm can not be used since it would register the atlas priors to the wrong segmentation solution. Figure 3 shows the log-likelihood curves: a lower likelihood indicates a better fit of the prior (atlas) to the posterior (classification) and therefore a better registration.

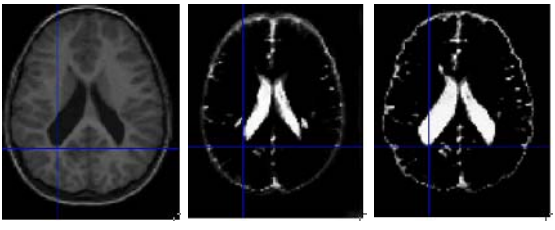


**Fig. 2.** PVL image (left); affine (middle) and EEM based segmentations (right)

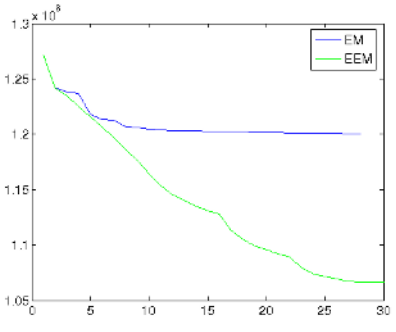



**Fig. 3.** Log-likelihood curves (vs. number of iterations) for experiment 2

## 4 Discussion

In [8] we present a comparison between 3 information theory based similarity measures. As discussed there, the CC measure, minimizing the statistical

distance between the apriori probabilities (the atlas) and aposteriori classification of the study image, is the best performing measure (according to the chosen evaluation parameters) among the 3. The CC measure has to rely on the aposteriori classification issued from the EM algorithm and this can originate errors when morphologically abnormal brains are involved. In order to circumvent this problem the segmentation algorithm and the CC non-rigid registration should cooperate in one algorithm.

In this paper we present a theoretic framework where atlas-to-study image non-rigid registration and image segmentation are merged. Both processes (non-rigid registration and segmentation) minimize the same functional, i.e. the log-likelihood function $Q$. The integration framework is very general and independent from the particular non-rigid registration regularizer. In fact we derive a dense force field as $\frac{\partial Q}{\partial T}$ that can be used as input for an entire class of regularizers (diffusion, elastic, curvature etc.).

In [1] a similar method is presented. Here we choose for a regularizer allowing larger deformations to be recovered. Therefore a different optimization was also used to compute the deformation field (gradient descent). Furthermore, we present here an extended validation on real images (having a ground truth) of normals and of patients (with no ground truth).

We performed two experiments. In the first experiment 20 T1 images, with known GM, WM and CSF (from manual segmentation) were used. The images were segmented by using different algorithms and the results were compared to the ground truth. It turned out that the EEM algorithm performed better (especially for CSF) than the different segmentation/registration schemes (EM-A and EM-II). The overlap coefficients related to the EEM methods were the highest when computed between the ground truth and the deformed atlas, indicating therefore a better registration. Better overlap coefficients (again, especially for CSF) were also achieved when considering the segmentation maps obtained with the different methods (EM-A, EM-II and EEM) indicating the solving registration and segmentation simultaneously (EEM) is better the solving them apart as done in EM-A and EM-II.

In the second experiment a PVL brain image was segmented twice, the first time using the EM algorithm (initialized by the affine registered atlas) and the second time using the EEM algorithm. This image presents highly deformed ventricles. As it can be seen in figure 2 the EEM outperforms the EM algorithm (since we don't have any ground truth here, only visual inspection is used).

In the future we will focus on the extension to more than 4 classes (WM, GM, CSF and OTHER) of the EEM algorithm. In this way we can make our algorithm more specific for a particular application where one particular brain structure has to be segmented and somehow quantified (for instance the hyppocampus in schizofrenia). Furthermore different regularizers can be integrated in our framework depending on the particular application.

## References


[1] J. Ashburner and K.J. Friston. Unified segmentation. *NeuroImage*, 26(2005), 839-851, 2005

[2] A. Hammers, M.J. Koep, S.L. Free, M. Brett, M.P. Richardson, C. Labbé, V.J. Cunningham, D.J. Brooks, and J. Duncan. Implementation and application of a brain template for multiple volumes of interest. *Human Brain Mapping*, 15(3):165–174, 2002.

[3] Q. Wang, D. Seghers, E. D'Agostino, F. Maes, D. Vandermeulen, P. Suetens and A. Hammers Construction and validation of mean shape atlas templates for atlas-based brain image segmentation in *Proceedings information processing in medical imaging (IPMI 2005),* volume 3565 of Lecture notes in computer science, pp. 689-700, July 2005, Glenwood Springs, Colorado, USA

[4] X. Chen, M. Brady, and D. Rueckert. Simultaneous segmentation and registration for medical image. In C. Barillot and D. Haynor, editors, *MICCAI 2004*, volume 3216 of *Lecture Notes in Computer Science*, pages 663–670, Saint-Malo, France, September 2004. Springer-Verlag, Berlin.

[5] E. D'Agostino, , F. Maes, D. Vandermeulen, B. Fischer, and P. Suetens. An information theoretic approach for non-rigid image registration using voxel class probabilities. In J.C. Gee, J.B.A. Maintz, and M.W. Vannier, editors, *Proc. Second International Workshop on Biomedical Image Registration (WBIR 2003)*, volume 2717 of *Lecture Notes in Computer Science*, pages 122–131, Philadelphia, PE, June 2003. Springer-Verlag, Berlin.

[6] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[7] E. D'Agostino,, F. Maes and D. Vandermeulen and P. Suetens, A viscous fluid model for multimodal non-rigid image registration using mutual information. *Medical Image Analysis*, 7(4):565–575, 2003

[8] E. D'Agostino,, F. Maes and D. Vandermeulen and P. Suetens, An information theoretic approach for non-rigid image registration using voxel class probabilities. *Medical Image Analysis* Accepted for publication, 2005

[9] Rohlfing, D. B. Russakoff, M. J. Murphy, and C.R. Maurer Jr. An intensity-based registration algorithm for probabilistic images and its application to 2d-3d image registration. In *in Proc. of SPIE: Medical Imaging 2002, San Diego, CA*, pages 581–591, San Diego, California, 2002.

[10] S. Smith, P. Bannister, C. Beckmann, M. Brady, S. Clare, D. Flitney, P. Hansen, M. Jenkinson, D. Leibovici, B. Ripley, M Woolrich, , and Y. Zhang. Fsl: New tools for functional and structural brain image analysis. *Seventh Int. Conf. on Functional Mapping of the Human Brain*, 2001.

[11] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of MR images of the brain. *IEEE Transactions on Medical Imaging*, 18(10):897–908, 1999.

[12] S. K. Warfield, M. Knaus, F. A. Jolesz, and R. Kikinis. Adaptive, template moderate, spatially varying statistical classification. *Medical Image Analysis*, 4(1):43–55, 2000.

[13] P. Wyatt and J. Alison Noble. Map mrf joint segmentation and registration of medical images. *Medical Image Analysis*, 7(4):539–552, 2003.

# Deformable Physiological Atlas-Based Programming of Deep Brain Stimulators: A Feasibility Study

Pierre-François D'Haese [1, 3], Srivatsan Pallavaram[1], Hong Yu[2], John Spooner[2], Peter E. Konrad[2], and Benoit M. Dawant[1]

[1] Department of Electrical Engineering and Computer Science, Vanderbilt University
[2] Department of Neurological Surgery, Vanderbilt University, Nashville, TN, 37235, USA
[3] Department of Electrical Engineering and Computer Science, Université Catholique de Louvain (UCL), 1348 Louvain-la-Neuve, Belgium

**Abstract.** The postoperative neurological management of patients with deep brain stimulation (DBS) of the subthalamic nucleus (STN) for Parkinson's disease is a complex and dynamic process that involves optimizing the stimulation parameters and decreasing the anti-parkinsonian medication while assessing the interactions of both treatment modalities. Neurologists who manage patients undergoing DBS therapy must have expert knowledge of the electro-anatomy of the subthalamic area and be familiar with the medical treatment of motor and non-motor symptoms. In clinical practice, finding the optimal programming parameters can be a challenging and time-consuming process. We have developed a computerized system to facilitate one of the bottlenecks of DBS therapy: the IPG (Internal Pulse Generator) programming. This system consists of a deformable physiological atlas built on more than 300 intra-operative macro-stimulations acquired from 30 Parkinson's patients and of a non-rigid registration algorithm used to map these data into an atlas. By correlating the position of the quadripolar electrode implanted in the patient with the information contained in our atlas, we can determine which of four contacts has the highest probability to be the most clinically effective. Preliminary results presented in this study suggest that this approach facilitates the programming process by guiding the neurologist to the optimal contact. The system we propose was tested retrospectively on a total of 30 electrodes. In 19 of these cases, this system predicted the contact that was selected as the optimal one by the neurologist.

## 1 Introduction

Since its first FDA approval in 1998, deep brain stimulation of the subthalamic nucleus (STN-DBS) has been established as an effective therapy for patients suffering from movement disorders [3], [4]. The therapy has significant applications in the treatment of tremor, rigidity, and drug induced side effects in Parkinson's disease (PD). Generally, the neurologist conducts the majority of DBS programming starting ~2 weeks after implantation. This allows the patient to recover from surgery and provides enough time for the transient lesional effects to resolve. Detailed principles

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 144 – 150, 2006.

and methods used to select the optimal programming parameters have been presented by different authors [1], [2].

Briefly, the first step in postoperative programming is the examination of the effectiveness and side effects induced by each individual contact. The electrode contacts are sequentially evaluated in a monopolar configuration in an effort to determine the contact that produces the best compromise. Frequency and pulse width are typically kept at constant settings of 130-180 Hz and 60-120 μs respectively. Amplitude is steadily increased to the tolerance level of the patient or until side effects occur. Repeated motor evaluation is then performed to assess the efficacy of stimulation. Ten to 15 minutes are allowed to pass between trials of separate contacts to allow the effects from previous stimulations to disappear. If a satisfactory result cannot be achieved with monopolar stimulation, more complex arrays consisting of bipole, tripole, or multiple cathodes are tried. The initial programming session, as described above, can take several hours and requires continuous feedback from the patient to ascertain the degree of benefit and to identify any side effects. This can be very taxing, especially when patients are kept off of medication for long periods of time. Furthermore, finding the optimal settings may take several trials over many months, which can be frustrating.

Automated selection of the optimal contact would facilitate the programming process and reduce the length of time required to determine optimum programming and thus be beneficial to the patients. In this paper, we propose a mechanism to do so. It consists of mapping the position of each of the contacts onto a statistical atlas, which assigns to each of the contacts a probability value for the contact to be the optimal one. This method requires several key ingredients: (1) accurate algorithms to register patients to the atlas, and (2) populating the atlas with data that permits the computation of the aforementioned probability. In our current system, the data we use is the response of previously implanted patients to intraoperative stimulations. In the remainder of this paper, we describe the method we have used as well as promising preliminary results.

## 2 Patients and Methods

Thirty PD patients undergoing DBS therapy have been enrolled in this study. With IRB approval (Vanderbilt University IRB #010809), a set of CT and MRI volumes were acquired pre-operatively for each patient. These were acquired with the patient anesthetized and head secured to the table to minimize motion. Typical CT images are acquired at kvp = 120 V, exposure = 350 mas, 512 x 512 voxels ranging in size from 0.49 to 0.62 mm, and slice thickness from 1 mm to 2 mm; MR images acquired with a 1.5T GE Signa scanner are 3D SPGR volumes, TR: 12.2, TE: 2.4, dimension 256x256x124 voxels, typical voxels dimensions 0.85x0.85x1.3 $mm^3$.

The surgical procedure as well as pre- and post-operative evaluations were identi-cal for all 30 patients. Seventeen of these were followed for a period of at least 6 months after DBS implantation and had optimal programming parameters determined by their neurologist or neurosurgeon. At the time of writing, the remaining 13 patients have not had long enough follow-up to achieve stable programming.

Surgical planning as well as the operative procedure performed at our institution has been described in detail in our previous work [6]. Briefly, pre-operative target

identification is performed automatically using an atlas-based method; automatically predicted targets are then checked by the functional neurosurgeon. This location is then refined intra-operatively based on the surgical team's interpretation of electrophysiological recordings and responses to stimulations; this team includes a neurosurgeon, a neurophysiologist, and a neurologist.

At our institution the procedure is performed with a miniature stereotactic frame, the StarFix microTargeting Platform® (501(K), Number K003776, Feb. 23, 2001, FHC, INC; Bowdoinham, ME) instead of a standard stereotactic frame. During surgery, a micropositioning drive (microTargeting® drive system, FHC Inc., Bowdoinham, ME) is mounted on the platform. Recording and stimulating leads are then inserted through the guiding tubes. The StarFix platform is designed based on the CT images (geometric distortions that affect the markers in MR images reduce platform accuracy when this modality is used) and its design is such that the pre-operative target is located on the central track. Details on the platform, including a study of its accuracy that shows it to be at least as accurate as standard frames can be found in [7]. The depth of the electrode is read from the micropositioning device and converted into x, y, and z CT coordinates. The x, y, and z position of each contact is computed using the geometry of the lead and the final intraoperative position of the center of the implant in CT coordinates. The implants used for these patients are the Medtronic 3389 implants, where the size of each contact is 0.5 mm and the gaps between the contacts are 0.5 mm.

### 2.1 Rigid and Non-rigid Registration Algorithms

A key component of the method we propose is our ability to map information acquired from a population of patients onto one reference image volume, which we call the atlas. Two types of registrations algorithms are needed to achieve this goal: rigid and non-rigid. The rigid registration algorithm is required to register MR and CT volumes of the same patient. This is necessary because, as mentioned above, the intra-operative positions of the electrode contacts provided by the micropositioning drive are in CT coordinates. The algorithm we have used to register MR and CT images of the same patient is an independent implementation of a standard Mutual Information-based algorithm [6]. Non-rigid registration is required to register patient data to the atlas and vice-versa. In this study, non-rigid registration is always performed on MR image volumes using an algorithm we have proposed recently [5]. Briefly, this algorithm computes a deformation field that is modeled as a linear combination of radial basis functions with finite support. The similarity measure we use is the Mutual Information between the images. We also compute simultaneously two transformations (one from the atlas to the subject and the other from the subject to the atlas) that are constrained to be inverses of each other.

While validation of non-rigid registration algorithms is an open-ended problem, in [6] we demonstrate our ability to register accurately MR volumes for STN-DBS implantation tasks. This, in turn, indicates that we are able to register accurately the patient volumes to the atlas.

### 2.2 Intra-operative Efficacy Atlas

Intra-operatively, macro-stimulation is performed to determine the optimal implant position. While targeting the STN, stimulation is applied approximately every 2 mm along the track, starting at the boundary of the STN, which is determined by micro-electrode recordings (MERs) acquired prior to stimulation. At every position, stimulation is typically performed with voltages starting at 0.5 V up to 5 V by 0.5 V increments. The effect of the stimulation on rigidity, muscle tone, bradykinesia, paresthesias, muscle contraction, eye movements and subjective sensations are assessed for every voltage. The optimal voltage is determined at each position and the loss of rigidity expressed in percent is recorded for this voltage. Because we can map the intra-operative coordinates of a patient's electrode onto the atlas, any information acquired intra-operatively can be projected onto the atlas. This, in turn, permits the creation of a number of statistical maps relating spatial coordinates in the atlas to characteristics measured intra-operatively. In [6] we have, for instance, shown that it is possible to create maps of features extracted from MERs. This study showed that maps of the mean spike rate can be used to define the boundary of the STN in the atlas. In this work, we have focused on developing maps that can provide useful information to the neurologist for programming. The key idea is to create an atlas that associates position with the efficacy of each electrode contact. Here, we define efficacy as being (1) proportional to the percent of loss of rigidity; (2) proportional to the therapeutic window, which equals the difference in voltage required to achieve this loss of rigidity (V) and the voltage at which side effects occur (VSE); and (3) inversely proportional to V. A position is good if the percent of loss of rigidity is high, V is low, and the difference between VSE and V is large. To create an atlas that captures this information, we first map the intra-operative stimulation position onto the atlas. At each position, we then center a Gaussian curve defined as follows:

$$F(x,y,z) = Loss\_of\_Rigidity * (V^{SE} - V) * \frac{1}{V} \exp\{-(\frac{x^2 + y^2 + z^2}{2V^2})\}$$

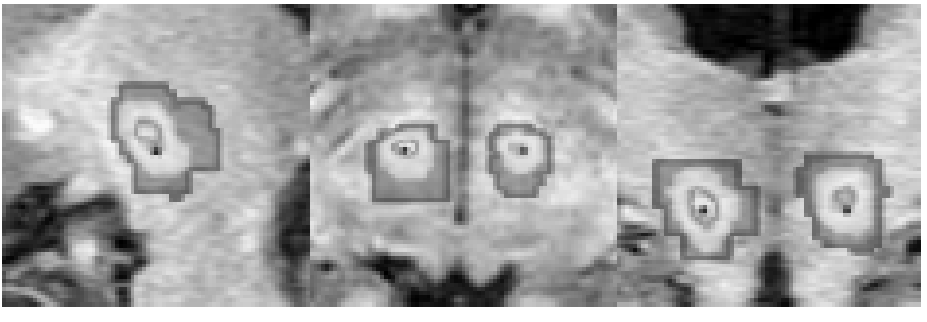


**Fig. 1.** Physiological stimulation map. White values represent a high likelihood to get good stimulation results, dark gray represents low likelihood to get good stimulation results. The star represents the optimal point in the atlas at which to place the implant when targeting the STN [6].

A point with a small stimulation voltage and a high loss of rigidity (in percent) will thus be associated with a curve with a small standard deviation and large amplitude. We repeat this procedure for every point for which we have intra-operative information and we produce a statistical atlas by averaging all these curves. In this atlas, a point associated with a curve that has a small standard deviation and large amplitude has a large but localized effect on the atlas. A point with a large standard deviation has a smaller impact that extends over a larger region. Fig. 1 illustrates results obtained with this method. In this figure, white means a high probability of obtaining good efficacy while dark gray means low probability of obtaining good efficacy. The black star is the average intra-operative position of the centers of all electrodes mapped onto the atlas for each side.

### 2.3 Atlas-Based Contact Selection

Once the atlas is created it can be used post-operatively to assist the neurologist in selecting the optimum contact for stimulation. To achieve this, the position of the patient's electrode is first mapped onto the atlas. The contact that falls into the area on the atlas corresponding to the highest probability of good efficacy would be the optimum contact for stimulation.

## 3 Results and Discussion

Table 1 shows quantitative results we have obtained with the method we've developed. We correlated the efficacy probability from our atlas to each contact in the 17 subjects included in the study. In table 1, the numbers in gray are the contacts selected by the neurologist. Contacts are numbered from C0 (distal contact) to C3 (proximal contact). The column labeled "V" is the amplitude of the therapeutic voltage.

Results show that about 60% of the contacts selected by the neurologist are the contacts with the highest efficacy probability in our atlas. Albeit preliminary, this supports the feasibility of using a statistical atlas to facilitate the programming process. A more detailed analysis of this process also suggests that using predictions from our atlas may shorten the time required to reach stable programming. For example, programming notes from the neurologists for patient P3 show that contact C1 was tried first on the right side before moving to C2 which produced better results. For patient P11 the C0 contacts were first tried on both sides before moving to contacts C2. A similar trend has been observed for the left implant in patient P12. Here, the neurologist moved from contact C3 to C2. For patient P15, contact C1 on the left side was observed to have a better effect on rigidity than contact C2.

For a few cases, the optimal electrode predicted by our atlas has been tried and rejected. In patient P16, contacts C0 and C1 were tried but not selected because these caused significant side effects. These effects were reduced with contact C3 but this particular patient still has significant rigidity and bradykinesia. For a number of cases, the optimal electrode predicted by our atlas has not been tried or programming records have not been available. Therefore, whether or not the electrode our atlas predicts would have led to better results cannot be determined.

**Table 1.** Shows, for 17 STN patients, the likelihood of the four contacts to produce good stimulation results. The number in gray shows the contact that was selected as the best one by the neurologist. Contacts are numbered from C0 to C3 (bottom to top contact). V shows the therapeutic voltage that was used.

| | Left | | | | | Right | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Patient | C0 | C1 | C2 | C3 | V | C0 | C1 | C2 | C3 | V |
| P0 | 0.47 | 0.77 | 0.84 | 0.61 | 1.1 | NI | 0.1 | 0.2 | 0.3 | 1.1 |
| P1 | 0.18 | 0.53 | 0.55 | 0.71 | 2.2 | 0.2 | 0.4 | 0.6 | 0.7 | 2.2 |
| P2 | 0.53 | 0.84 | 0.94 | 0.61 | 2.4 | 0.4 | 0.7 | 0.9 | 1.0 | 1.7 |
| P3 | 0.64 | 0.78 | 0.69 | 0.27 | 1.5 | 0.2 | 0.4 | 0.7 | 1.0 | 1.3 |
| P4 | 0.06 | 0.07 | 0.10 | 0.24 | 1.5 | 0.3 | 0.6 | 0.7 | 0.4 | 1.8 |
| P5 | 0.01 | 0.11 | 0.12 | 0.30 | 1.8 | | | | | |
| P6 | 0.74 | 0.74 | 0.69 | 0.47 | 1.5 | 0.4 | 0.7 | 0.8 | 0.7 | 1.5 |
| P7 | 0.65 | 0.59 | 0.47 | 0.21 | 1.6 | 0.7 | 0.9 | 1.0 | 0.9 | 1.7 |
| P8 | 0.80 | 0.47 | 0.23 | 0.10 | 1.4 | | | | | |
| P9 | 0.73 | 0.98 | 0.93 | 0.65 | 2.3 | 0.8 | 0.9 | 0.8 | 0.4 | 2.2 |
| P10 | 0.72 | 0.84 | 0.84 | 0.59 | 3.2 | 0.3 | 0.6 | 0.6 | 0.4 | 3.2 |
| P11 | 0.59 | 0.85 | 0.90 | 0.65 | 1.8 | 0.2 | 0.4 | 0.6 | 0.9 | 1.5 |
| P12 | 0.60 | 0.49 | 0.41 | 0.11 | 2.1 | 0.2 | 0.4 | 0.6 | 0.5 | 2.7 |
| P13 | | | | | | 0.9 | 1.0 | 0.9 | 0.8 | 2.3 |
| P14 | 0.12 | 0.47 | 0.84 | 0.84 | 0.8 | NI | 0.1 | 0.2 | 0.4 | 1.0 |
| P15 | 0.73 | 0.91 | 0.80 | 0.40 | 1.5 | 0.3 | 0.5 | 0.7 | 0.6 | 1.5 |
| P16 | 0.85 | 0.92 | 0.69 | 0.50 | 1.8 | 0.3 | 0.6 | 0.7 | 0.6 | 1.8 |
| P17 | | | | | | 0.2 | 0.4 | 0.7 | 1.0 | 3.2 |

The results presented in this study demonstrate that a computer-assisted method can be developed to facilitate what remains a bottleneck in DBS therapy. A number of improvements on the method presented herein are currently being developed. First, a prospective validation study has been initiated. Rather than verifying that the electrode we propose is the optimal one after programming has been completed, we will propose the optimal contact to the neurologist at the time of initial programming. We have followed this approach when developing and validating our automatic pre-operative target prediction for DBS implantation [6]. Second, at the time of programming, we will provide the neurologist with a 3D display of the position of the electrodes in the efficacy map overlaid on high resolution MR images. This will permit correlation of these positions with anatomy, thereby facilitating spatial orientation and navigation between the contacts. Third, as the number of patients increases, we will create maps of side effects. Currently, we only use a crude definition of efficacy: reduction in rigidity weighted by

the therapeutic voltage window (i.e., the difference between the voltage required to suppress the symptoms and voltage inducing side effects). We will refine this definition to improve the way side effects are taken into consideration. To achieve this, we will create maps of side effects as we have done for our current definition of efficacy. This will permit an automatic multi-parameter optimization procedure that will minimize side effects while maximizing the positive effects of the stimulation.

## Acknowledgements

This project was supported, in parts, by a Vanderbilt University Discovery Grant. Pierre-François D'Haese is supported in parts by the FRIA/FNRS (Belgian Science Foundation).

## References

[1] Pollak P, Krack P, Fraix V, et al. Intraoperative micro- and macrostimulation of the STN in Parkinson's disease. Mov Dis., 2002; 17(Suppl. 3):S155–S161

[2] Volkmann J, Herzog J, Kopper F, Deuschl G. Introduction to the programming of deep brain stimulation. Mov Dis. 2002; 17, S181–S187

[3] R. G. Deuschl, J. Volkmann, and P. Krack, "Deep brain stimulation for movement disorders," Mov Dis., 2002 vol. 17, pp. S1–S11.

[4] B. Schrader, W. Hamel, D. Weinert, and H. M. Mehdorn, "Documentation of electrode localization." Mov Dis, vol. 17 pp. S167–S174, 2002.

[5] G. K. Rohde, A. Aldroubi, and B. M. Dawant, "The adaptive bases algorithm for intensity based nonrigid image registration," registration," IEEE Transactions on Medical Imaging, vol. 22, pp. 1470–1479, 2003.

[6] P.F. D'Haese, E. Cetinkaya, P.E. Konrad, C. Kao, B.M. Dawant, "Computer-aided placement of deep brain stimulators: from planning to intraoperative guidance" IEEE Transactions on Medical Imaging, vol. 24 (11), pp. 1469-78, Nov 2005.

[7] Fitzpatrick JM, Konrad PE, Nickele Ch, Cetinkaya E, and Kao Ch: "Accuracy of Customized Miniature Stereotactic Platforms", Stereotactic and Functional Neurosurgery 2005; 83:25-31

# A Comparison of Acceleration Techniques for Nonrigid Medical Image Registration

Stefan Klein, Marius Staring, and Josien P.W. Pluim

University Medical Center Utrecht, Image Sciences Institute,
Q0S.459, P.O. Box 85500, 3508 GA Utrecht, The Netherlands
{stefan, marius, josien}@isi.uu.nl.

**Abstract.** Mutual information based nonrigid registration of medical images is a popular approach. The coordinate mapping that relates the two images is found in an iterative optimisation procedure. In every iteration a computationally expensive evaluation of the mutual information's derivative is required. In this work two acceleration strategies are compared. The first technique aims at reducing the number of iterations, and, consequently, the number of derivative evaluations. The second technique reduces the computational costs *per iteration* by employing stochastic approximations of the derivatives. The performance of both methods is tested on an artificial registration problem, where the ground truth is known, and on a clinical problem involving low-dose CT scans and large deformations. The experiments show that the stochastic approximation approach is superior in terms of speed and robustness. However, more accurate solutions are obtained with the first technique.

## 1 Introduction

Nonrigid registration is an important technique in medical image processing. A popular class of registration methods is based on maximisation of the mutual information similarity measure, in combination with a deformation field parameterised by cubic B-splines [1,2]. The large computation time of this approach is a big disadvantage for many clinical applications. For practical use, acceleration is required.

Registration is usually stated as a minimisation problem:

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \mathcal{C}\left(\boldsymbol{\mu}; I_F, I_M\right) \quad , \tag{1}$$

where $\mathcal{C}(\boldsymbol{\mu}; I_F, I_M)$ denotes a cost function, and $\boldsymbol{\mu}$ a vector of parameters defining the deformation field that relates the fixed image $I_F$ and the moving image $I_M$. In this paper the cost function is defined as the negated mutual information similarity metric, and the deformation field is parameterised by cubic B-splines, whose coefficients form the vector $\boldsymbol{\mu}$. To find the solution $\hat{\boldsymbol{\mu}}$ an iterative optimisation strategy is employed. In many optimisation methods the parameter update in each iteration $k$ is based on the derivative of the cost function $\mathcal{C}$ with respect to the parameters $\boldsymbol{\mu}$. The gradient descent algorithm [3] is the most straightforward example:

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 151–159, 2006.

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - a_k \boldsymbol{g}(\boldsymbol{\mu}_k), \quad k = 0, 1, 2, \ldots \ , \tag{2}$$

where $\boldsymbol{g}(\boldsymbol{\mu}_k)$ represents the derivative of the cost function, $\partial \mathcal{C}/\partial \boldsymbol{\mu}$, evaluated at $\boldsymbol{\mu}_k$. A scalar gain, $a_k$, controls the step size. Within certain conditions, the sequence $\{\boldsymbol{\mu}_k\}$ defined by (2) converges to a local minimum of the cost function.

Computation of the derivatives, $\boldsymbol{g}(\boldsymbol{\mu}_k)$, requires a considerable amount of computational effort in nonrigid registration problems. In this work two acceleration strategies are compared. The first technique aims at reducing the number of iterations, and, consequently, the number of required derivative evaluations. Well-known methods with an improved rate of convergence are the quasi-Newton and nonlinear conjugate gradient [3]. The second technique focusses on the computational costs *per iteration* by using approximations of $\boldsymbol{g}(\boldsymbol{\mu}_k)$. Acceleration factors are given with respect to the performance of the standard gradient descent method.

The acceleration strategies are compared in two types of experiments. Firstly, an artificially created problem is considered. An image is registered to itself, after application of a known, randomly generated deformation. Secondly, the registration of a low-dose expiration CT chest scan to a high-dose inspiration scan of the same patient is used as a test problem.

## 2 Nonrigid Registration Framework

In this section the various components of the nonrigid registration framework are described. The design of our algorithm is largely based on the papers by Mattes et al. [1], Rueckert et al. [2], and Thévenaz and Unser [4].

For computation of the mutual information the approach described in [4] is used. The joint histogram is constructed using B-spline Parzen windows, which makes it possible to formulate the mutual information as a continuous, differentiable function of the parameters describing the deformation field. In all experiments described in this paper, the joint histogram size is set to $32 \times 32$. The deformation field is parameterised by B-splines.

The minimisation problem (1) is solved with a multiresolution strategy. For the image data, we use a Gaussian image pyramid. The complexity of the deformation model is defined by the B-spline control point resolution. In our tests we let it follow the image resolution: when the image resolution is doubled, the control point resolution is doubled as well. The number of resolution levels and the final B-spline control point spacing depend on the specific problem. At each resolution a minimisation is performed, using one of the tested optimisation methods. When necessary, images are rigidly registered before the nonrigid registration.

## 3 Acceleration Strategies

A standard gradient descent algorithm, see (2), is used as a benchmark, to which further results are compared. The method is implemented using a slowly decaying gain sequence:

$$a_k = a/(A + k + 1)^\alpha \ , \tag{3}$$

where $a > 0$, $A \geq 0$, and $0 \leq \alpha \leq 1$ are user-defined constants.

### 3.1 Acceleration by Faster Convergence

In the literature many optimisation methods can be found with (theoretically) a better rate of convergence than the gradient descent. Perhaps the most well-known ones are the quasi-Newton and nonlinear conjugate gradient methods.

Quasi-Newton methods [3] use the following iterative scheme:

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - a_k L_k \boldsymbol{g}(\boldsymbol{\mu}_k) \ . \tag{4}$$

In this equation, $L_k$ is a positive definite matrix that serves as an approximation to the inverse Hessian of the cost function. For computation of the matrix $L_k$ second order derivatives of the cost function are *not* needed; only the already computed first order derivatives are used. The scalar $a_k$ is again a gain factor that controls the step size.

Several ways to construct the series $\{L_k\}$ are proposed in the literature. The method used in this work is a popular variant of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm: the Limited memory BFGS (LBFGS) [5], which eliminates the need for storing the matrix $L_k$ in memory.

Quasi-Newton methods are usually implemented in combination with an inexact line search routine, which determines a gain factor $a_k$ that ensures sufficient progress towards the solution. In this work we use the line search routine described by Moré and Thuente [6]. If no gain factor can be found that gives sufficient progress, the optimisation is assumed to have converged.

Nonlinear conjugate gradient methods [3, 7] are based on the following iterative scheme:

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k + a_k \boldsymbol{d}_k \ , \tag{5}$$

where the search direction $\boldsymbol{d}_k$ is defined as a linear combination of the current derivative $\boldsymbol{g}(\boldsymbol{\mu}_k)$ and the previous search direction $\boldsymbol{d}_{k-1}$:

$$\boldsymbol{d}_k = -\boldsymbol{g}(\boldsymbol{\mu}_k) + \beta_k \boldsymbol{d}_{k-1} \ . \tag{6}$$

Many expressions for the scalar $\beta_k$ can be found in the literature [7]. In this study we use a so-called 'hybrid' version, proposed in [8] and shown to be very efficient compared to other methods. The gain factor $a_k$ is determined by the same inexact line search routine as used with the quasi-Newton method.

### 3.2 Acceleration by Stochastic Approximation

By using approximated derivatives instead of the exact ones the computation time per iteration can be reduced significantly.

The computation time of the derivative of mutual information is linearly dependent on the number of voxels $|I_F|$ in the fixed image, and on the number of B-spline coefficients $N$ (the length of the parameter vector $\boldsymbol{\mu}$):

$$t_{\boldsymbol{g}(\boldsymbol{\mu}_k)} \sim p|I_F| + qN \ , \tag{7}$$

where $p$ and $q$ are constants. For most medical nonrigid registration problems $p|I_F|$ tends to be much larger than $qN$. It is clear that we can lower the computation time of a derivative evaluation by not using all the voxels, but only a small subset of voxels.

In [9] it is demonstrated that, when using a new, randomly selected, subset of voxels in every iteration of the optimisation process, the parameter sequence $\{\boldsymbol{\mu}_k\}$ still converges to the correct solution. Selecting a new subset of voxels in every iteration ensures that the approximation errors will, on average, cancel each other out. The approximation errors can be considered a source of noise $\boldsymbol{\varepsilon}_k$ entering the optimisation process:

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k - a_k \left( \boldsymbol{g}(\boldsymbol{\mu}_k) + \boldsymbol{\varepsilon}_k \right) \quad . \tag{8}$$

This scheme is often referred to as a *stochastic gradient descent algorithm* or a Robbins-Monro procedure [10, 11]. It can be proven that the sequence $\{\boldsymbol{\mu}_k\}$ defined by (8) still converges to the solution $\hat{\boldsymbol{\mu}}$, provided that the *bias* of the approximation error goes to zero.

The experiments in [9] indicate that for the registration of large 3D images as few as 2048 voxels are required in each iteration, which is adopted in our tests. The gain sequence $\{a_k\}$ is defined as in the gradient descent method, see (3).

### 3.3 Combining the Acceleration Strategies

Naturally, the question rises whether it is possible to combine the two acceleration strategies. Unfortunately, the quasi-Newton and nonlinear conjugate gradient optimisation methods are not designed to work with stochastic approximations of the derivatives. They expect noise-free derivatives to be available.

A possible strategy for these methods is to select *a single* subset of voxels in the fixed image and use these samples throughout the registration process [12,1]. A disadvantage of this method is that convergence to the correct solution cannot be guaranteed, because the approximation error bias does not go to zero.

In our tests with quasi-Newton and nonlinear conjugate gradient the samples are selected on a regular grid using identical downsampling factors for each image dimension.

## 4 Experiments and Results

In two types of nonrigid registration problems, the following methods are compared:

- Gradient Descent (GD), see Sec. 3,
- Quasi-Newton (QN-*df*), see Sec. 3.1,
- Nonlinear Conjugate Gradient (NCG-*df*), see Sec. 3.1,
- Stochastic Gradient Descent (SGD), see Sec. 3.2.

The extension *df* denotes the downsampling factor. Downsampling factors of 1 (full image), 2, 4, 8, and 16 are tested. In case of GD, the full image is always

used. As explained in Sec. 3.2, the SGD method uses only 2048 voxels to calculate the derivatives.

For all optimisation methods, the computation time per iteration is assumed to be dominated by the time required for computing the derivative of the mutual information. Moreover, the derivative's computation time is assumed to be mainly related to the number of voxels used. These simplifications allow us to define the *normalised computation time (NCT)* up to iteration $k$ of the optimisation process:

$$\mathrm{NCT} = (k+1)V/|I_F| \ , \tag{9}$$

with $V$ the number of voxels used to compute the derivative.

### 4.1 Artificial Deformation Fields

In the first evaluation procedure an image $I$ is registered with a deformed version of itself. To avoid interpolation errors, the deformation field $\tilde{\boldsymbol{v}}$ is added to the B-spline deformation field $\boldsymbol{v}_{\boldsymbol{\mu}}$ that is updated during optimisation. Since the image $I$ is registered with itself, the desired solution is a total deformation field that is zero everywhere. The ground truth is known, so an error measure, the average displacement error $e$, can be defined:

$$e(\boldsymbol{\mu}) = \frac{1}{|I|} \sum_{\boldsymbol{x}_i \in I} \| \tilde{\boldsymbol{v}}(\boldsymbol{x}_i) + \boldsymbol{v}_{\boldsymbol{\mu}}(\boldsymbol{x}_i) \| \ , \tag{10}$$

where $\boldsymbol{x}_i$ is the position of voxel $i$ in the image volume $I$, and $|I|$ the total number of voxels in $I$. The speed of convergence of a method is visualised by plotting this error measure against the normalised computation time NCT.

The experiments are performed on four 3D CT images of the heart. The images originate from chest scans. These were cropped to the area of the heart and downsampled by a factor of two, resulting in images of $97 \times 97 \times 97$ voxels with an isotropic size of 1.4 mm. For each image a deformation field $\tilde{\boldsymbol{v}}$ is generated, composed of randomly placed Gaussian blobs with a standard deviation of 14 mm. A $10 \times 10 \times 10$ grid of B-spline control points defines the deformation field $\boldsymbol{v}_{\boldsymbol{\mu}}$, yielding 3000 parameters to be optimised. No multiresolution schemes are used in this experiment, which makes comparison of the results more straightforward. The maximum number of iterations is limited to 2048. Three constants must be set for the gain sequence in (3): $a = 3200$, $A = 50$, and $\alpha = 0.602$.

In the following we present the test results for one of the four images. The outcome for the other images is similar. Figure 1 shows the average displacement error as a function of the normalised computation time for the tested methods. A logarithmic scale is used for the time axis. It is clear that both acceleration strategies realise considerable speed improvements compared to the standard GD procedure. The QN and NCG method do not give acceptable results anymore with a downsampling factor of 8, which results in 2197 selected voxels. The methods QN-16 and NCG-16, which are omitted from the figure, perform even worse. Without downsampling QN and NCG achieve a slightly higher accuracy

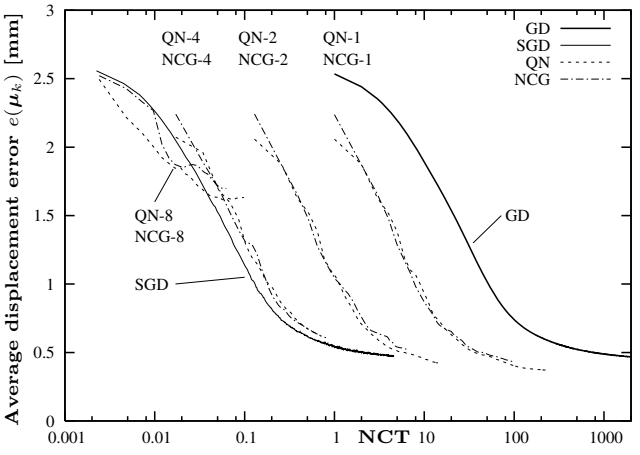



**Fig. 1.** Results for one of the heart images. The graph shows the average displacement error as a function of the normalised computation time for the tested methods.

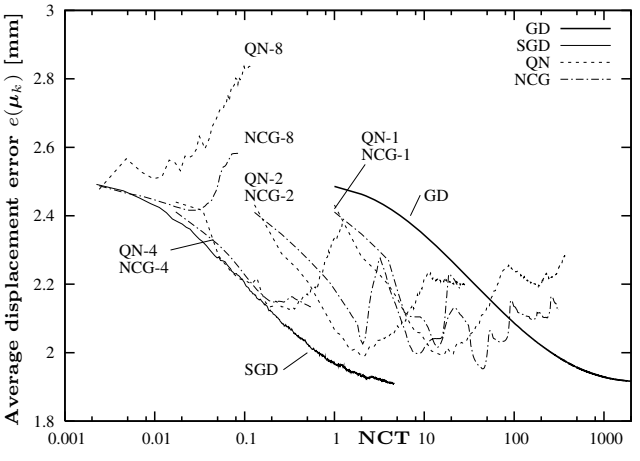



**Fig. 2.** The imposed deformation field is composed of Gaussian blobs with a standard deviation of 7 mm. The graph shows the average displacement error as a function of the normalised computation time for the tested methods.

than the GD and SGD method. The SGD method, which works well with only 2048 voxels, is clearly the fastest.

The tests are repeated for a more difficult registration problem. The imposed deformation field $\tilde{\boldsymbol{v}}$ is composed of Gaussian blobs with a standard deviation of 7 mm. This smaller standard deviation results in a deformation field that is very

hard to recover, since the B-spline control point grid used during registration is not dense enough. The test results for the same image as before are shown in Fig. 2. It is interesting to see that the QN and NCG methods can not handle this very ill-posed registration problem. The GD and SGD procedures remain stable. As expected, none of the optimisation methods is able to achieve a very large reduction of the initial average displacement error, since the B-spline control point grid is not fine enough.

Note that the QN and NCG methods *do* still find a set of parameters that decrease the mutual information. The Moré-Thuente line search, employed in both QN and NCG to set the gain factor $a_k$, guarantees that the cost function decreases in every iteration, $\mathcal{C}(\boldsymbol{\mu}_k) < \mathcal{C}(\boldsymbol{\mu}_{k-1})$. However, in an ill-posed problem a decreasing cost function does not imply that actual progress is made towards the correct solution.

### 4.2 Clinical Data

In this section a number of experiments with 3D CT chest scans are described. The patients were scanned after inspiration and after expiration. The inspiration scans were recorded with a high radiation dose; the expiration scans with a low dose. The large deformations in combination with the noisy nature of the low-dose scans make this a challenging registration problem.

The images were acquired with a Philips Mx8000IDT 16-slice CT scanner. We use data of seven patients. The original images, with in-plane dimensions of $512 \times 512$ and a number of slices ranging from 400 to 800, were downsampled by a factor of 2 in each dimension to be able to register the images on a standard PC with one gigabyte of memory. The resulting voxel size is approximately 1.4 mm in all directions, and the images consist of about $10^7$ voxels.

A four-level multiresolution approach is applied. At each resolution the number of iterations is fixed to 256. At the highest resolution the B-spline control point spacing is set to 22 mm, yielding a grid of about $19^3$ control points (approximately 20000 parameters). The following settings are used for the gain sequence in GD and SGD: $a = 60000$, $A = 50$, and $\alpha = 0.602$.

As is common in clinical applications of nonrigid registration, the ground truth is not known. We assess the registration results by computing the overlap of the lungs, $L_1$ and $L_2$:

$$\text{Overlap} = \frac{2 \cdot |L_1 \cap L_2|}{|L_1| + |L_2|} \quad . \tag{11}$$

Segmentations of the lungs were made by means of a method based on the work of Hu et al. [13]. In the segmentations large pulmonary vessels are not considered part of the lungs. For reference: a translation of one voxel in each dimension results in a lung overlap of about 0.95.

Table 1 shows the overlap measures after rigid registration and nonrigid registration with the studied methods. Each column displays the results for a single

**Table 1.** The results of CT chest scan registration for patients P1-P7. The overlap of segmented lungs is shown after rigid registration (first row), and after nonrigid registration with different methods. Each column contains the results for a single patient.

| | P1 | P2 | P3 | P4 | P5 | P6 | P7 |
|---|---|---|---|---|---|---|---|
| rigid | 0.736 | 0.453 | 0.729 | 0.883 | 0.654 | 0.811 | 0.671 |
| GD | 0.925 | 0.802 | 0.951 | 0.972 | 0.951 | 0.975 | 0.914 |
| SGD | 0.922 | 0.799 | 0.945 | 0.966 | 0.945 | 0.972 | 0.909 |
| QN-1 | 0.932 | 0.855 | 0.956 | 0.974 | 0.962 | 0.979 | 0.930 |
| QN-2 | 0.931 | 0.836 | 0.955 | 0.973 | 0.960 | 0.979 | 0.926 |
| QN-4 | 0.926 | 0.815 | 0.951 | 0.971 | 0.955 | 0.975 | 0.916 |
| QN-8 | 0.905 | 0.797 | 0.934 | 0.960 | 0.936 | 0.965 | 0.898 |
| QN-16 | 0.862 | 0.751 | 0.906 | 0.938 | 0.897 | 0.946 | 0.858 |
| NCG-1 | 0.931 | 0.795 | 0.941 | 0.966 | 0.942 | 0.969 | 0.929 |
| NCG-2 | 0.912 | 0.796 | 0.955 | 0.973 | 0.961 | 0.969 | 0.904 |
| NCG-4 | 0.911 | 0.792 | 0.940 | 0.966 | 0.943 | 0.969 | 0.903 |
| NCG-8 | 0.895 | 0.794 | 0.935 | 0.959 | 0.934 | 0.965 | 0.894 |
| NCG-16 | 0.854 | 0.752 | 0.903 | 0.938 | 0.900 | 0.946 | 0.858 |

patient. All methods result in a considerable improvement on the rigid registration. The overlap measures confirm the results found in Sec. 4.1. The final accuracy of the nonrigid registration is, compared to GD, very little affected by the random subsampling strategy employed by SGD. The methods QN-1, QN-2, NCG-1, and NCG-2 result in a somewhat better accuracy than GD and SGD. With higher downsampling factors the accuracy decreases.

The most remarkable results are found in patient P2. The QN-1 and QN-2 methods seems to outperform all other methods. However, visual inspection of the results taught us that the good overlap results come at the price of some very unrealistic deformations. In patients P5 and P7 the same problem was observed for the QN and NCG methods. The GD and SGD procedures only have this problem in patient P5. This is in line with the results of Sec. 4.1, where GD and SGD also seem to be more robust than QN and NCG.

## 5 Conclusion

We have compared acceleration techniques for both artificially deformed and real clinical data. The experiments indicate that SGD achieves the largest acceleration, and seems to be more robust for badly defined problems than the QN and NCG algorithms. Without downsampling QN and NCG yield a slightly smaller error than GD or SGD. Downsampling increases the error, and does not result in the same acceleration as obtained by SGD.

In summary, we can conclude that the acceleration technique focussing on reduction of the computational costs per iteration is the preferred approach.

## Acknowledgements

Funding of this research has been provided by the Netherlands Organisation for Scientific Research (NWO). Additionally, this work benefited from the use of the Insight Segmentation and Registration Toolkit (ITK), an open source software developed as an initiative of the U.S. National Library of Medicine and available at `http://www.itk.org`.

## References

1. Mattes, D., Haynor, D.R., Vesselle, H., Lewellen, T.K., Eubank, W.: PET-CT image registration in the chest using free-form deformations. IEEE Transactions on Medical Imaging **22** (2003) 120–128
2. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: Application to breast MR images. IEEE Transactions on Medical Imaging **18** (1999) 712–721
3. Nocedal, J., Wright, S.J.: Numerical optimization. Springer-Verlag, New York (1999)
4. Thévenaz, P., Unser, M.: Optimization of mutual information for multiresolution image registration. IEEE Transactions on Image Processing **9** (2000) 2083–2099
5. Nocedal, J.: Updating quasi-Newton matrices with limited storage. Mathematics of Computation **35** (1980) 773–782
6. Moré, J.J., Thuente, D.J.: Line search algorithms with guaranteed sufficient decrease. ACM Transactions on Mathematical Software **20** (1994) 286–307
7. Dai, Y.H.: A family of hybrid conjugate gradient methods for unconstrained optimization. Mathematics of Computation **72** (2003) 1317–1328
8. Dai, Y.H.: An efficient hybrid conjugate gradient method for unconstrained optimization. Annals of Operations Research **103** (2001) 33–47
9. Klein, S., Staring, M., Pluim, J.P.W.: Comparison of gradient approximation techniques for optimisation of mutual information in nonrigid registration. In Fitzpatrick, J., Reinhardt, J., eds.: SPIE Medical Imaging: Image Processing. Volume 5747 of Proceedings of SPIE., SPIE Press, Bellingham, WA (2005) 192–203
10. Robbins, H., Monro, S.: A stochastic approximation method. The Annals of Mathematical Statistics **22** (1951) 400–407
11. Kushner, H.J., Clark, D.S.: Stochastic approximation methods for constrained and unconstrained systems. Springer-Verlag, New York (1978)
12. Kybic, J., Unser, M.: Fast parametric elastic image registration. IEEE Transactions on Image Processing **12** (2003) 1427–1442
13. Hu, S., Hoffman, E.A., Reinhardt, J.M.: Automatic lung segmentation for accurate quantitation of volumetric X-Ray CT images. IEEE Transactions on Medical Imaging **20** (2001) 490–498

# Evaluation of Similarity Measures for Non-rigid Registration

Darko Škerl, Boštjan Likar, and Franjo Pernuš

University of Ljubljana, Faculty of Electrical Engineering
Tržaška 25, 1000 Ljubljana, Slovenia
{darko.skerl, bostjan.likar, franjo.pernus}@fe.uni-lj.si

**Abstract.** In this paper we propose a protocol for the evaluation of similarity measures for non-rigid registration. The protocol is feasible for the evaluation of non-rigid registration in which the deformation model is based on a set of regularly or irregularly distributed corresponding point pairs. The proposed protocol is able to deduce five properties of the similarity measure for each point pair of the deformation model, so that local or global estimation of the similarity measure properties can be derived. The feasibility of the proposed protocol was demonstrated on a popular deformation model based on B-splines, on six similarity measures, and on the "gold standard" CT and MR images of three spine vertebrae and three MR T1 and T2 images of the head.

## 1 Introduction

Medical image registration has emerged as a particularly active field due to many clinical applications. A general overview of registration techniques is given in [1, 2]. Image registration is concerned with finding a spatial transformation that will bring two images into spatial correspondence. According to the nature of geometrical transformation, registration techniques can be divided into rigid and non-rigid ones. A rigid transformation is composed of rotations and translations, while non-rigid transformation can be modeled by the spline warps, truncated basis function expansions, Navier-Lamé equations, or by a viscous fluid model. Non-rigid registration is of great importance for integration of complimentary information of mono-modal temporal images or multimodal images of the same anatomy. Other possible applications of non-rigid registration are inter-subject registration, by which anatomical variations within a certain population can be studied, or atlas to patient registration by which arbitrary information from the atlas can be transferred to the patient.

The accuracy and robustness of a registration method depends on the imaging modality, image content, image quality, spatial transformation, similarity measure, optimization, and numerous implementation details. The complex interdependence of these factors makes the assessment of the influence of a particular factor on registration difficult. Consequently, several methods for validation of registration methods have been proposed [3-6]. One of the major obstacles in validation of registration methods is a lack of true anatomic correspondence between the images, called "gold standard", which can usually be established only for a limited set of image pairs. This

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 160 – 168, 2006.

problem is especially apparent in the evaluation of non-rigid registration for which a precise correspondence between all image points cannot be established for real images. To overcome this problem numerical simulations of the deformations have been proposed [6].

The similarity measure is one of the factors that most affect the quality of registrations. Traditionally, limited information on the behavior of a similarity measure is obtained either by studying the quality of the final registration or by drawing plots of similarity measure as a function of spatial transformation parameters. In our previous publication [7] we have presented a protocol for a more thorough and optimization-independent evaluation of similarity measures for rigid registration (http://lit.fe.uni-lj.si/Evaluation).

In this paper we extend the protocol to evaluate the similarity measures for non-rigid registrations. We show that the protocol is feasible for the evaluation of similarity measures that are based on a set of regularly or irregularly distributed corresponding point pairs, for example, when spatial deformations are modeled by spline warps or by truncated basis functions.

## 2 Methods and Materials

### 2.1 Evaluation Protocol

To evaluate the similarity measure for non-rigid registrations, we derive several properties that characterize the behavior of similarity measures for each pair of corresponding pairs of the spatial deformation model. For this purpose a "gold standard" registration between the floating and the target image is needed. The similarity measure is a function of all the parameters of the deformation model. In 3D, each point pair can be displaced in $x$, $y$ or $z$ directions so that the number of parameters of the deformation model is three times larger than the number of corresponding point pairs. However, to derive the properties of the similarity measure for a single point pair, all the other point pairs can be fixed, so that the similarity measure can be looked upon as a function of the displacement of only one pair of corresponding points.

Let the "gold standard" position of the analyzed point pair define the origin $\mathbf{X}_0$ of the 3-dimensional parametrical space ($x$, $y$ and $z$) and let SM($\mathbf{X}$) be the value of a similarity measure at point $\mathbf{X}$; $\mathbf{X}=[x,y,z]$ in this space. Similarity measure values SM($\mathbf{X}_{n,m}$) are defined along $N$ lines and at $M+1$ points evenly spaced on each of the $N$ lines, defined by a randomly selected starting position $\mathbf{X}_{n,-M/2}$ at a distance $R$ from the origin $\mathbf{X}_0$ and its mirror point $\mathbf{X}_{n,M/2}$; $\mathbf{X}_{n,-M/2}=-\mathbf{X}_{n,M/2}$. The number of lines $N$ were determined experimentally [7] so that the properties of the similarity measures did not change if $N$ was increased. Before computing the similarity measure profiles along each of the $N$ lines and for each point pair, all the other point pairs can be randomly displaced from the "gold standard" position for a predefined distance $RD$ by which a more realistic simulation of the registration process can be achieved. All similarity measure values are normalized to the interval [0, 1] according to the minimal and

maximal values of the similarity measure before normalization. For each point pair the properties of the similarity measures are computed from the obtained similarity measure profiles along the *N* lines:

1. Accuracy (*ACC*) of a similarity measure is defined as the root mean square (RMS) of distances between the origin $\mathbf{X}_0$ and the positions $\mathbf{X}_{n,opt}$, n=1, 2, …, *N* along the *N* lines, where the SM(**X**) reaches an optimum.
2. Distinctiveness of optimum (*DO*) is the estimation of the uncertainty of the location of the optimum for a given point pair.
3. Capture range (*CR*) is defined as the smallest of the *N* distances between the optima and the closest minima.
4. Number of minima (*NOM(r)*) is the sum of minima of the similarity measure within distance *r* from each of the *N* global optima, i.e. a cumulative number of minima as a function of distance *r*.
5. Risk of nonconvergence (*RON*) is the property that describes the number and extent of a similarity measure around the *N* global optima.

Details on the computation of the similarity measure properties can be found in [7] and (http://lit.fe.uni-lj.si/Evaluation). The first two properties describe the behavior of similarity measure close to the "gold standard", while the last three properties estimate the robustness of a similarity measure.

### 2.2 Images

We have used the MR and CT images of three spine vertebrae L2, L3 and L4 [8] and three MR T1 and T2 images of the head[1] for which "gold standard" rigid registrations were known. Fig. 1 shows 2D slices of some of the 3D images used in the experiments. For the vertebrae the rigid "gold standard" transformation was obtained using fiducial markers while for the brain images stereotactic frames were used to compute the best rigid match.

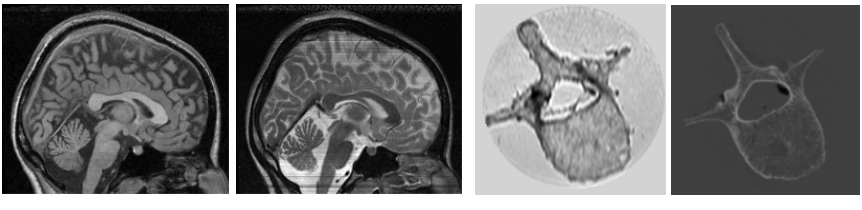


**Fig. 1.** T1 and T2 slices of the 3D image of head (left) and MR and CT slices of the 3D image of the L3 vertebra (right)

### 2.3 Deformation Model

Among numerous existing spatial deformation models, based on a set of regularly or irregularly distributed corresponding points, we have chosen one of the most

[1] http://www.loni.ucla.edu/ICBM/

frequently used ones, i.e. the model based on B-splines interpolation [9] which was proposed for non-rigid registration by Rueckert et. al.[10]. The deformation model is defined on a set of regularly spaced point pairs, termed knots. The number of knots defines the flexibility, while the degree of the B-splines defines the smoothness of the interpolation.

In our experiments we have used 3rd degree B-splines and a 3D grid of 7×7×7 knots. The evaluation of the five properties of the similarity measures was performed on the 5×5×5 central knots (375 parameters of the deformation model), while the knots on the edges were fixed. Before computing the similarity measure profile along each of the *N* lines for each knot all the other knots were randomly displaced from the "gold standard" position for a distance *RD* = 1 mm by which a more realistic simulation of the registration process was achieved (Fig. 2).

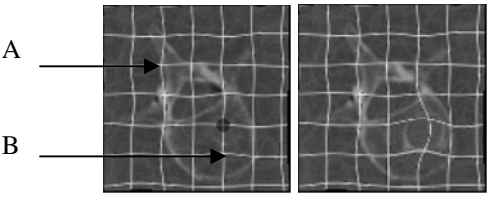



**Fig. 2.** Left: deformed image in all knots for *RD* = 1 mm (left), except for the knot in a dark circle, which was fixed. Right: the same image with the knot that was fixed in the left image displaced for 20 mm. The points (A, B) illustrate the knots for which results are given in Table 1.

## 3 Results

In the results section we show how the mutual information (MI) [11, 12], normalized mutual information (NMI) [13], recently proposed asymmetric gradient based multi feature mutual information (AMMI) [14], correlation ratio [15], entropy (H) and energy of the histogram (E) similarity measures behave when different parts of vertebra and brain are deformed. Partial volume interpolation [11] was used to compute the joint histogram for all six similarity measures. The parameters of the evaluation protocol *R, N* and *M* were set to 20 mm, 50 and 80, respectively.

Fig. 3 shows interpolated values of the five similarity measure properties of MI for the vertebrae L2-L4. The values of the properties are printed in the legend to the right of each image. For all three vertebrae, mutual information is more accurate and robust in the spinous process and lamina area than in the area of the vertebra bodies.

Fig. 4 shows five properties for non-rigid registration of three sets of real MR T1 and MR T2 images. It can be seen that mutual information is less accurate and robust in the parietal lobe area. The robustness is also poor in the frontal lobe area. The results among different images are consistent.

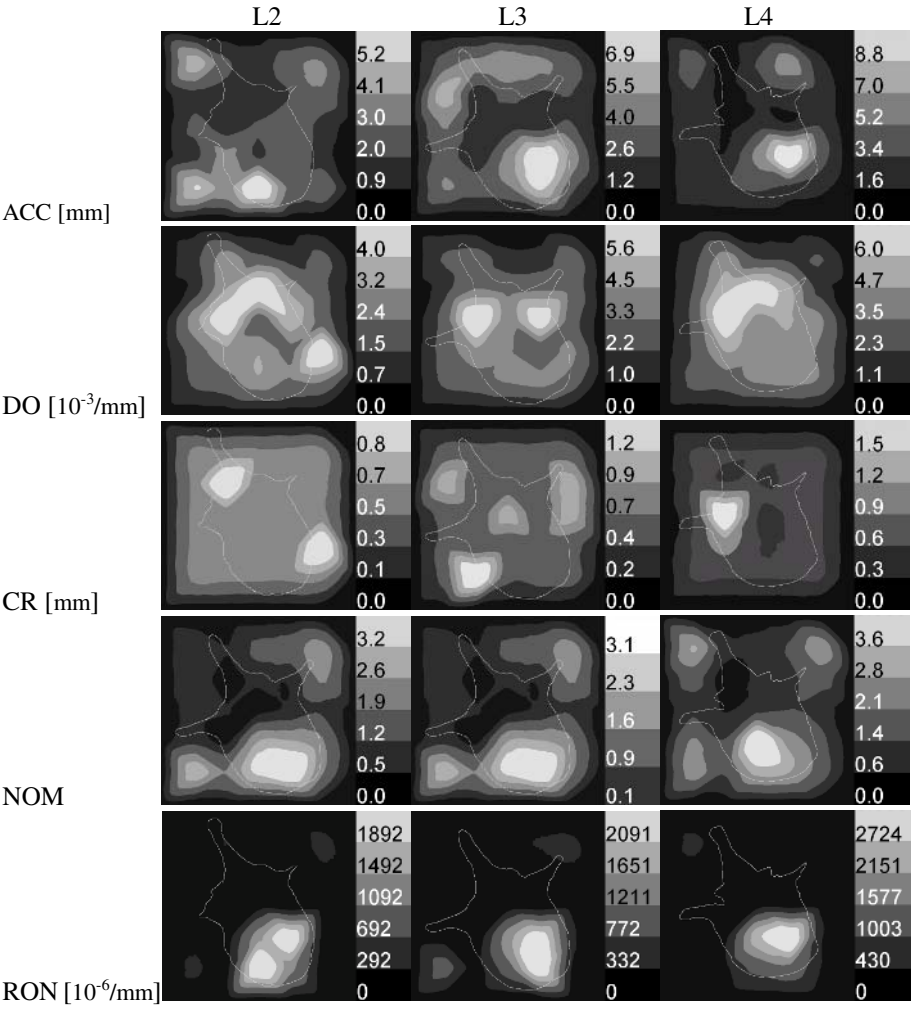



**Fig. 3.** Five properties (from top to bottom) of MI for non-rigid registration of MR and CT images of the L2 (left), L3 (middle) and L4 (right) vertebrae. Contours the vertebrae are overlaid on each image.

Table 1 shows quantitative results of the evaluation of the 6 similarity measures for non-rigid registration of three sets of MR and CT images of the three vertebrae. Local comparison of the similarity measures is given for the points A and B, while average similarity measure properties for all points provide a global comparison of similarity measures. Both MI and NMI are accurate and robust around the spinous process and lamina (point A), i.e. where AMMI is also robust but less accurate. In the body area (point B) MI and NMI are less accurate, while AMMI is far more accurate but less robust. Global comparisons in which the average and standard deviations of the five

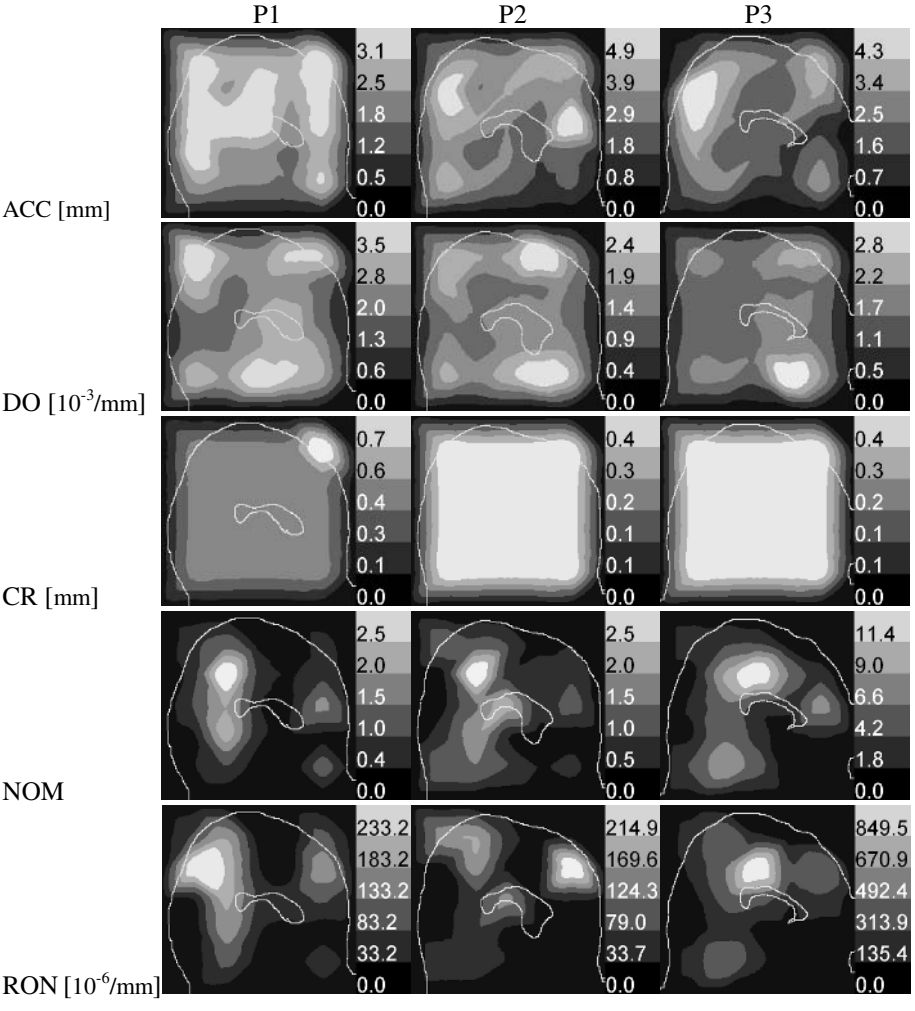



**Fig. 4.** Five properties (from top to bottom) of MI for non-rigid registration of three pairs of MR T1 and MR T2 images of the head. A corresponding contour of the skull and corpus callosum is overlaid on each image

properties for all points and images, given at the bottom of Table 1, show that MI and NMI have very similar properties. In general, AMMI is more accurate but less robust similarity measure than MI and NMI. COR is the least accurate out of all analyzed measures.

Global comparisons of the similarity measure properties for the three sets of MR T1 and T2 images of the brain are given in Table 2. Also for the brain images, we can conclude that AMMI is more accurate but less robust similarity measure than MI and NMI. Again, MI and NMI have very similar values of the properties, which are much better than of COR, H and E.

**Table 1.** The properties of the three most accurate similarity measures for non-rigid registration of three sets of MR and CT images of the vertebrae. The point A (see Fig. 2) is where MI was accurate and robust while the point B is where MI was less accurate and robust.

| | MI | | | NMI | | | AMMI | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Local comparison: Point A** | | | | | | | | |
| | L2 | L3 | L4 | L2 | L3 | L4 | L2 | L3 | L4 |
| ACC | 1.2 | 1.4 | 0.6 | 1.1 | 1.4 | 0.5 | 3.2 | 1.4 | 0.8 |
| DO | 3.0 | 3.1 | 5.2 | 2.8 | 3.1 | 5.4 | 6.3 | 6.4 | 11.8 |
| CR | 1.0 | 0.5 | 0.4 | 1.0 | 0.9 | 7.0 | 0.5 | 0.5 | 0.4 |
| NOM | 0.1 | 0.1 | 0.1 | 0.1 | 0.0 | 0.1 | 3.6 | 6.8 | 10.9 |
| RON | 2.0 | 6.1 | 6.1 | 2.8 | 4.3 | 7.2 | 585.3 | 1071.5 | 1447.6 |
| | **Local comparison: Point B** | | | | | | | | |
| ACC | 2.5 | 6.7 | 2.2 | 2.3 | 6.0 | 1.9 | 0.6 | 0.9 | 0.5 |
| DO | 2.1 | 3.4 | 3.5 | 2.4 | 3.0 | 3.8 | 10.2 | 10.0 | 15.1 |
| CR | 0.5 | 0.5 | 0.4 | 0.5 | 0.5 | 0.4 | 0.5 | 0.5 | 0.4 |
| NOM | 3.4 | 3.8 | 2.7 | 3.6 | 4.3 | 2.6 | 21.7 | 19.1 | 23.5 |
| RON | 1139.2 | 2376.3 | 1302.2 | 1093.5 | 1901.8 | 1129.4 | 3106.6 | 2826.0 | 4057.2 |

| | **Global comparisons: Average values (std) of the five properties** | | | | | |
|---|---|---|---|---|---|---|
| | MI | NMI | AMMI | COR | H | E |
| ACC | 2.7 (1.7) | 2.6 (1.7) | 1.2 (0.7) | 4.5 (2.7) | 3.5 (1.8) | 3.3 (1.4) |
| DO | 2.8 (1.4) | 2.8 (1.3) | 9.4 (2.9) | 1.9 (1.5) | 1.5 (0.6) | 1.2 (0.7) |
| CR | 0.5 (0.2) | 0.7 (0.8) | 0.5 (0.0) | 0.5 (0.1) | 0.5 (0.0) | 0.5 (0.0) |
| NOM | 1.4 (1.1) | 1.4 (1.1) | 14.5 (6.3) | 2.9 (1.6) | 2.4 (1.2) | 4.4 (1.7) |
| RON | 388 (686) | 326 (554) | 2020 (1166) | 641 (795) | 228 (247) | 324 (259) |

**Table 2.** The properties of the six similarity measures for non-rigid registration of three sets of MR T1 and MR T2 images of the brain

| | **Global comparisons: Average values (std) of the five properties** | | | | | |
|---|---|---|---|---|---|---|
| | MI | NMI | AMMI | COR | H | E |
| ACC | 2.8 (1.0) | 2.8 (1.0) | 0.9 (0.5) | 7.7 (4.1) | 4.2 (2.3) | 6.5 (3.3) |
| DO | 1.9 (0.7) | 1.8 (0.7) | 7.6 (2.4) | 2.3 (2.3) | 0.9 (0.4) | 0.5 (0.3) |
| CR | 0.4 (0.1) | 0.4 (0.1) | 0.4 (0.0) | 0.5 (0.2) | 0.4 (0.1) | 0.4 (0.1) |
| NOM | 1.6 (2.3) | 1.7 (2.2) | 15.0 (7.0) | 1.7 (1.1) | 2.8 (2.2) | 3.0 (2.1) |
| RON | 107 (157) | 105 (151) | 1565 (1054) | 1095 (1546) | 220 (319) | 210 (314) |

## 4 Conclusion

We have extended the protocol for evaluation of similarity measures for rigid registration [7] to the evaluation of non-rigid registration, which is based on a set of regularly or irregularly distributed corresponding point pairs. With the proposed protocol we are able to deduce five properties of the similarity measure for each point pair of the deformation model, so that local estimation of the similarity measure properties can be derived. The feasibility of the proposed protocol was demonstrated on the "gold standard" CT and MR images of three spine vertebrae and three MR T1 and T2 images of the head. Non-rigid registration is often needed for this kind of images

because, for example, brain can deform by up to 20 mm after the skull is opened during neurosurgery [16] or because MR images can be geometrically deformed due to the inhomogeneity of the magnetic field.

The proposed protocol enables quantitative local and global estimation of the similarity measure properties for non-rigid registration. The results of such evaluations can help the researchers to select the best similarity measure and the appropriate optimization scheme for a given registration task. The protocol may also be useful for studying the best combination of similarity measures by which robust and accurate non-rigid registration can be achieved. For example, the results of this study suggest that registration process should start with MI or NMI and continue with AMMI similarity measure.

## References

[1] P. M. Thompson and A. W. Toga, "Warping strategies for intersubject registration," in Handbook of Medical Image Processing, I. Bankman, Ed.: Academic Press, 1999, pp. 569-601.

[2] J. B. Maintz and M. A. Viergever, "A survey of medical image registration," Med Image Anal, vol. 2, pp. 1-36, 1998.

[3] R. P. Woods, "Validation of Registration Accuracy," in Handbook of Medical Image Processing, I. Bankman, Ed.: Academic Press, 1999, pp. 491-498.

[4] P. Hellier, C. Barillot, I. Corouge, B. 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 Trans Med Imaging, vol. 22, pp. 1120-30, 2003.

[5] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer, Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. Hill, D. J. Hawkes, C. Studholme, J. B. Maintz, M. A. Viergever, G. Malandain, R. P. Woods, and et al., "Comparison and evaluation of retrospective intermodality brain image registration techniques," J Comput Assist Tomogr, vol. 21, pp. 554-66, 1997.

[6] J. A. Schnabel, C. Tanner, A. D. Castellano-Smith, A. Degenhard, M. O. Leach, D. R. Hose, D. L. Hill, and D. J. Hawkes, "Validation of nonrigid image registration using finite-element methods: application to breast MR images," IEEE Trans Med Imaging, vol. 22, pp. 238-47, 2003.

[7] D. Škerl, B. Likar, and F. Pernuš, "A protocol for evaluation of image registration similarity measures," in press: IEEE Trans Med Imaging, 2006.

[8] D. Tomaževič, B. Likar, T. Slivnik, and F. Pernuš, "3-D/2-D registration of CT and MR to X-ray images," IEEE Trans Med Imaging, vol. 22, pp. 1407-1416, 2003.

[9] M. Unser, "Splines: A Perfect Fit for Signal and Image Processing," IEEE Signal Processing Magazine, vol. 16, pp. 22-38, 1999.

[10] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," IEEE Trans Med Imaging, vol. 18, pp. 712-21, 1999.

[11] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," IEEE Trans Med Imaging, vol. 16, pp. 187-98, 1997.

[12] W. M. Wells, 3rd, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," Med Image Anal, vol. 1, pp. 35-51, 1996.

[13] C. Studholme, D. L. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," Pattern Recognition, vol. 32, pp. 71-86, 1999.
[14] D. Tomaževič, B. Likar, and F. Pernuš, "3-D/2-D registration by integrating 2-D information in 3-D," IEEE Trans Med Imaging, vol. 25, pp. 17-27, 2006.
[15] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodality image registration," in Medical Image Computing and Computer Assisted Intervention, vol. 1496: Springer Verlag, 1998, pp. 1115-1124.
[16] T. Hartkens, D. L. Hill, A. D. Castellano-Smith, D. J. Hawkes, C. R. Maurer, Jr., A. J. Martin, W. A. Hall, H. Liu, and C. L. Truwit, "Measurement and analysis of brain deformation during neurosurgery," IEEE Trans Med Imaging, vol. 22, pp. 82-92, 2003.

# Computing the Geodesic Interpolating Spline

Anna Mills[1], Tony Shardlow[1,⋆], and Stephen Marsland[2,⋆⋆]

[1] The University of Manchester, UK
amills@ma.man.ac.uk, shardlow@ma.man.ac.uk
[2] Massey University, NZ
S.R.Marsland@massey.ac.nz

**Abstract.** We examine non-rigid image registration by knotpoint matbreak ching. We consider registering two images, each with a set of knotpoints marked, where one of the images is to be registered to the other by a nonlinear warp so that the knotpoints on the *template* image are exactly aligned with the corresponding knotpoints on the *reference* image. We explore two approaches for computing the registration by the Geodesic Interpolating Spline. First, we describe a method which exploits the structure of the problem in order to permit efficient optimization and second, we outline an approach using the framework of classical mechanics.

## 1 Introduction and Formulation of Problem

We develop methods for computation of the Geodesic Interpolating Spline (GIS) as described by Marsland and Twining in [1]. A comprehensive survey of registration methods is given in [2]. The GIS was developed from the thin-plate spline [3] and clamped-plate spline [4] to provide diffeomorphic mapping between images so that no folds or tears are introduced to the images and no information is lost from the image being mapped.

We formulate the GIS as a minimization problem as follows. Minimize

$$l(\boldsymbol{q}_i(t), \boldsymbol{v}(t,\boldsymbol{x})) = \frac{1}{2}\int_0^1 \int_B \|L\boldsymbol{v}(t,\boldsymbol{x})\|^2_{\mathbb{R}^d} \mathrm{d}\boldsymbol{x}\,\mathrm{d}t \ , \tag{1}$$

over deformation fields, $\boldsymbol{v}(t,\boldsymbol{x}) \in \mathbb{R}^d$ and paths, $\boldsymbol{q}_i(t) \in B$ for $i = 1,\ldots,n_c$, where $B \subset \mathbb{R}^d$ is the unit-ball domain of the image and $L = \nabla^2$ approximating the Willmore energy. We have constraints for $0 \le t \le 1,\ i = 1,\ldots,n_c$,

$$\frac{\mathrm{d}\boldsymbol{q}_i}{\mathrm{d}t} = \boldsymbol{v}(t,\boldsymbol{q}_i(t)), \boldsymbol{q}_i(0) = \boldsymbol{P}_i,\ \ \boldsymbol{q}_i(1) = \boldsymbol{Q}_i \ . \tag{2}$$

⋆ Partially supported by EPSRC grant GR/R78725/01.
⋆⋆ Partially supported by RSNZ Marsden Fund (Fast Start) grant MAU0408 "A principled approach to the non-rigid registration and structural analysis of groups of medical images".

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 169–177, 2006.

Using techniques in [5], we can represent the minimizing vector field as

$$\boldsymbol{v}(t,\boldsymbol{x}) = \sum_{i=1}^{n_c} \boldsymbol{\alpha}_i(t) G(\boldsymbol{q}_i(t),\boldsymbol{x}) \ , \tag{3}$$

where $\boldsymbol{\alpha}_i(t), \boldsymbol{q}_i(t)$ are respectively multipliers and knotpoint positions at time $t$ for a set of $n_c$ knotpoints and where $G(\boldsymbol{x},\boldsymbol{y})$ is the Green's function derived by Boggio [6]. We experiment numerically with the 2-dimensional Green's function $G(\boldsymbol{x},\boldsymbol{y}) = -|\boldsymbol{x}-\boldsymbol{y}|^2 \ln|\boldsymbol{x}-\boldsymbol{y}|^2$ for the biharmonic operator with zero Dirichlet and Neumann boundary conditions on $B$. In this way, we can derive

$$\min \int_0^1 \frac{1}{2} \sum_{i,j=1}^{n_c} \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_j G(\boldsymbol{q}_i,\boldsymbol{q}_j)\mathrm{d}t, \ \text{ such that } \ \frac{\mathrm{d}\boldsymbol{q}_i}{\mathrm{d}t} = \sum_{j=1}^{n_c} \boldsymbol{\alpha}_j G(\boldsymbol{q}_i,\boldsymbol{q}_j) \ , \tag{4}$$

$$\boldsymbol{q}_i(0) = \boldsymbol{P}_i, \qquad \boldsymbol{q}_i(1) = \boldsymbol{Q}_i, \qquad i = 1,\ldots,n_c \ . \tag{5}$$

We explore two methods for the minimization. In Sec. 2, we examine the structure of the discretized version of (4), and use an optimization method exploiting this structure. In Sec. 3, we reformulate the problem in a Hamiltonian framework to compute the GIS. In Sec. 4, we test the methods on brain images with knotpoints marked by clinicians. In Sec. 5, we summarize our results.

## 2 Numerical Optimization Exploiting Partial Separability

To find the minimizer of (4), we discretize in time to achieve a finite dimensional system. This yields a constrained optimization problem with a clear structure. The numerical optimization routine, Lancelot B, from the Galahad suite [7], uses group partial separable structure to express the dependence between different variables and make the structure of the resulting Hessian matrices clear, improving the performance on large scale problems. The discretization of the GIS is partially separable, because of the dependence on time. We move to a discretized version of the problem using time step $\Delta t = 1/N$. We use $\boldsymbol{\alpha}_i^n \approx \boldsymbol{\alpha}_i(n\Delta t), n = 0,\ldots,N-1$ and $\boldsymbol{q}_i^n \approx \boldsymbol{q}_i(n\Delta t), n = 0,\ldots,N$ to give the problem in the following form: Minimize

$$l\{\boldsymbol{\alpha}_i^n, \boldsymbol{q}_i^n\} = \frac{1}{2} \sum_{i,j=1}^{n_c} \sum_{n=0}^{N-1} \boldsymbol{\alpha}_i^{n\top} \boldsymbol{\alpha}_j^n G(\boldsymbol{q}_i^n, \boldsymbol{q}_j^n) \tag{6}$$

over $\boldsymbol{\alpha}_i^n, \boldsymbol{q}_i^n \ \ i = 1,\ldots,n_c$ such that

$$N(\boldsymbol{q}_i^{n+1} - \boldsymbol{q}_i^n) = \sum_{j=1}^{n_c} \boldsymbol{\alpha}_j^n G(\boldsymbol{q}_i^n, \boldsymbol{q}_j^n), \qquad n = 0,\ldots,N-1, \ i = 1,\ldots,n_c \tag{7}$$

with conditions $\boldsymbol{q}_i^0 = \boldsymbol{P_i}$, and $\boldsymbol{q}_i^N = \boldsymbol{Q}_i, \ i = 1,\ldots,n_c$. As detailed in [7], Lancelot B uses an iterative method. Outer iterations minimize augmented Lagrangian merit functions, and inner iterations minimize a quadratic model of

the merit function. Lancelot B minimizes a function where the objective function and constraints can be described in terms of a set of group functions and element functions, where each element function involves only a small subset of the minimization variables. Specifically, Lancelot B solves problems with objective functions of the form

$$f(\boldsymbol{x}) = \sum_{i \in \Gamma_0} w_i^g g_i \left( \sum_{j \in \mathcal{E}_i} w_{ij}^e e_j(\boldsymbol{x}_j^e) \right), \qquad \boldsymbol{x} = (x_1, x_2, \ldots, x_n)^\top \ . \tag{8}$$

In the above, we have $\Gamma_0$, a set of indices of group functions $g_i$ and we have $\mathcal{E}_i$, a set of nonlinear element functions $e_j$ and group and element weight parameters, respectively $w_i^g$ and $w_{ij}^e$. For the Lancelot implementation, the constraints must be of the form

$$c_i(\boldsymbol{x}) = w_i^g g_i \left( \sum_{j \in \mathcal{E}_i} w_{ij}^e e_j(\boldsymbol{x}_j^e) + \boldsymbol{a}_i^N \boldsymbol{x} \right) = 0 \ , \tag{9}$$

for $i$ in the set of indices of constraints $\Gamma_c$. Examining our objective function (6), we see that there is a natural division into $N$ groups, each group being given by, for the $n^{th}$ group

$$w_i^g g_i \left( \sum_{j \in \mathcal{E}_n} w_{nj}^e e_j(\boldsymbol{x}_j^e) \right) = \frac{1}{2} \sum_{i,j=1}^{n_c} (\boldsymbol{\alpha}_i^n G(\boldsymbol{q}_i^n, \boldsymbol{q}_j^n)) \boldsymbol{\alpha}_j^n \ . \tag{10}$$

In this notation, $\boldsymbol{x}_j^e$ is the vector containing the optimization variables $(\boldsymbol{q}_i^n, \boldsymbol{\alpha}_i^n)$, for $i = 1, \ldots, n_c$. Similarly, the $2n_cN$ velocity constraints in (7) can be characterized by $2n_cN$ groups. We require that the start and end points of each control point path coincide with the landmarks on, respectively, the floating and reference images. This gives $4n_c$ constraints of the form

$$0 = (\boldsymbol{q}_i^1 - \boldsymbol{P}_i)w, \qquad 0 = (\boldsymbol{q}_i^{N+1} - \boldsymbol{Q}_i)w, \qquad i = 1, \ldots, n_c \ . \tag{11}$$

Hence we have $4n_c$ groups characterizing the landmark constraints, each group being weighted by some $w \gg 1$. In total, we have $N + 2n_cN + 4n_c$ groups characterizing the problem.

We see the block diagonal sparsity structure of the Hessian for the augmented Lagrangian function for our problem in Fig. 1 where the Hessian for a problem involving 5 time steps and 8 knotpoints is shown, calculated using a numerical scheme on the augmented Lagrangian function as used in Lancelot B,

$$\mathcal{L}_A(\boldsymbol{x}, \boldsymbol{\lambda}; \mu) = f(\boldsymbol{x}) - \sum_{i \in \Gamma_c} \lambda_i c_i(\boldsymbol{x}) + \frac{1}{2\mu} \sum_{i \in \Gamma_c} c_i^2(\boldsymbol{x}) \ , \tag{12}$$

where $\mu$ is the penalty parameter and $\lambda_i$ for $i \in \Gamma_c$ are Lagrangian multipliers. There are 5 time blocks on the main diagonal, each with coupling to the adjacent time blocks. There is a banded off-diagonal structure, the two bands of

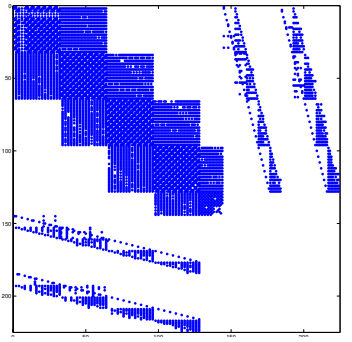


**Fig. 1.** The sparsity structure of the Hessian

$\mathcal{A}$ being due to the constraints being divided into $x$-dimension and $y$-dimension constraints. Lancelot B uses Newton methods, which require Hessian approximations. The routine exploits group partial separability of the problem to make calculation and storage of the Hessian approximations more efficient.

## 3 Classical Mechanics Approach

We present a novel formulation of the GIS for image registration as a problem in Hamiltonian dynamics. The GIS problem is given by the minimization problem (4). We can treat this as a Lagrangian by setting

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{1}{2} \sum_{i,j=1}^{n_c} \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_j G(\boldsymbol{q}_i, \boldsymbol{q}_j) \ , \tag{13}$$

where $\boldsymbol{q} = (\boldsymbol{q}_1, \ldots, \boldsymbol{q}_{n_c})$, and $\dot{\boldsymbol{q}} = (\frac{\mathrm{d}\boldsymbol{q}_1}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}\boldsymbol{q}_{n_c}}{\mathrm{d}t})$ represent position and velocity, respectively. Following Arnold [8], we see that the Hamiltonian of the system, $H(\boldsymbol{p}, \boldsymbol{q}) = \boldsymbol{p}^\top \dot{\boldsymbol{q}} - L(\boldsymbol{q}, \dot{\boldsymbol{q}})$, is the Legendre transform of the Lagrangian function as a function of the velocity, $\dot{\boldsymbol{q}}$, where $\boldsymbol{p}$ is the generalized momentum, $\frac{\partial L}{\partial \dot{q}}$, and so

$$\frac{\partial H}{\partial \boldsymbol{p}} = \dot{\boldsymbol{q}}, \qquad \frac{\partial H}{\partial \boldsymbol{q}} = -\frac{\partial L}{\partial \boldsymbol{q}} \ . \tag{14}$$

Construct a matrix $\mathcal{A} \in \mathbb{R}^{n_c \times n_c}$ so that $\mathcal{A}_{ij} = G(\boldsymbol{q}_i, \boldsymbol{q}_j),\ i, j = 1, \ldots, n_c$, and a matrix $\mathcal{G} \in \mathbb{R}^{dn_c \times dn_c}$ as $\mathcal{G} = \mathcal{A} \otimes I$, (where I is the d-dimensional identity matrix). We can use (4) and define a vector $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1^\top, \boldsymbol{\alpha}_2^\top, \ldots, \boldsymbol{\alpha}_{n_c}^\top]^\top$ to derive $\frac{\mathrm{d}\boldsymbol{q}}{\mathrm{d}t} = \mathcal{G}\boldsymbol{\alpha}$, and hence we see that the generalized momentum is given by $\frac{\partial L}{\partial \dot{\boldsymbol{q}}} = \mathcal{G}^{-1}\mathcal{G}\boldsymbol{\alpha} = \boldsymbol{\alpha}$. Substituting this expression into the standard Euler-Lagrange equations [8] gives $\partial L/\partial \boldsymbol{q} = \mathrm{d}\boldsymbol{\alpha}/\mathrm{d}t$. Hence, with (14), we have the coupled system of Hamiltonian equations

$$\dot{\boldsymbol{q}} = \frac{\partial H}{\partial \boldsymbol{\alpha}}, \qquad \dot{\boldsymbol{\alpha}} = -\frac{\partial H}{\partial \boldsymbol{q}}, \qquad \begin{bmatrix} \boldsymbol{q}(0) \\ \boldsymbol{\alpha}(0) \end{bmatrix} = \begin{bmatrix} \boldsymbol{P} \\ \boldsymbol{A} \end{bmatrix} = \boldsymbol{Y} \ , \tag{15}$$

where $\boldsymbol{P}$ is the vector of initial knotpoint positions in (4) and $\boldsymbol{A}$ is the initial vector of generalized momentum. Substituting (13) and the derivatives from (15) into the expression for the Hamiltonian of the system gives

$$H(\boldsymbol{\alpha}, \boldsymbol{q}) = \boldsymbol{\alpha}^\top \dot{\boldsymbol{q}} - L(\boldsymbol{q}, \dot{\boldsymbol{q}}) = \frac{1}{2} \sum_{i,j=1}^{n_c} \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_j G(\boldsymbol{q}_i, \boldsymbol{q}_j) \ , \tag{16}$$

so that $H$ is now a function of $\boldsymbol{\alpha}$ and $\boldsymbol{q}$. We have shown that the solutions, $\boldsymbol{q}_i(t)$ and $\boldsymbol{\alpha}_i(t)$ of (4) are solutions of (15). We solve the nonlinear system of equations $\Phi(\boldsymbol{A}; \boldsymbol{P}) = \boldsymbol{Q}$ for $\boldsymbol{A}$ as a shooting problem, where $\Phi(\boldsymbol{A}; \boldsymbol{P}) := \boldsymbol{q}(1)$, the position component of the solution of the Hamiltonian system

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{q}_i = \sum_{j=1}^{n_c} \boldsymbol{\alpha}_j G(\boldsymbol{q}_i, \boldsymbol{q}_j), \ \ \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\alpha}_i = -\sum_{j=1}^{n_c} \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_j \frac{\partial}{\partial \boldsymbol{q}_j} G(\boldsymbol{q}_i, \boldsymbol{q}_j), \ i = 1, \ldots, n_c \ , \tag{17}$$

with initial conditions given in (15).

*Numerical Implementation.* To solve (17), we discretize in time. We choose to discretize using the Forward Euler method. Experiments with symplectic methods have shown no advantage for this problem, principally because it is a boundary value problem where long time simulations are not of interest, and no suitable explicit symplectic integrators are available. Using the notation $\boldsymbol{q}_i^n \approx \boldsymbol{q}_i(n\Delta t)$, $\boldsymbol{\alpha}_i^n \approx \boldsymbol{\alpha}_i(n\Delta t), \ n = 0, \ldots, N, \ \Delta t = 1/N$, we have

$$\begin{bmatrix} \boldsymbol{q}^{n+1} \\ \boldsymbol{\alpha}^{n+1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}^n \\ \boldsymbol{\alpha}^n \end{bmatrix} + \Delta t \begin{bmatrix} \dfrac{H(\boldsymbol{q}^n, \boldsymbol{\alpha}^n)}{\partial \boldsymbol{\alpha}} \\ -\dfrac{H(\boldsymbol{q}^n, \boldsymbol{\alpha}^n)}{\partial \boldsymbol{q}} \end{bmatrix}, \begin{bmatrix} \boldsymbol{P} \\ \boldsymbol{A} \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}^0 \\ \boldsymbol{\alpha}^0 \end{bmatrix} = \boldsymbol{Y} \ . \tag{18}$$

We wish to examine the variation with respect to the initial momentum, $\boldsymbol{A}$ in order to provide Jacobians for the nonlinear solver. The initial positions, $\boldsymbol{P}$ remain fixed. Using the Forward Euler scheme for some function $\boldsymbol{f}$, we have $\boldsymbol{X}^{n+1} = \boldsymbol{X}^n + \Delta t \boldsymbol{f}(\boldsymbol{X}^n)$ with initial condition $\boldsymbol{X}^0 = \boldsymbol{Y}$. Differentiating with respect to $\boldsymbol{A}$ gives

$$\frac{\mathrm{d}\boldsymbol{X}^{n+1}}{\mathrm{d}\boldsymbol{A}} = \frac{\mathrm{d}\boldsymbol{X}^n}{\mathrm{d}\boldsymbol{A}} + \Delta t \frac{\mathrm{d}\boldsymbol{f}(\boldsymbol{X}^n)}{\mathrm{d}\boldsymbol{X}^n} \frac{\mathrm{d}\boldsymbol{X}^n}{\mathrm{d}\boldsymbol{A}}, \qquad \frac{\mathrm{d}\boldsymbol{X}^0}{\mathrm{d}\boldsymbol{A}} = [0, I]^\top \ , \tag{19}$$

where $I$ is the $2n_c \times 2n_c$ identity matrix. Let $J^n$ be $\frac{\mathrm{d}\boldsymbol{X}^n}{\mathrm{d}\boldsymbol{A}}$ and solve numerically a coupled system of equations

$$J^{n+1} = J^n + \Delta t \frac{\mathrm{d}\boldsymbol{f}(\boldsymbol{X}^n)}{\mathrm{d}\boldsymbol{X}^n} J^n, \qquad \boldsymbol{X}^{n+1} = \boldsymbol{X}^n + \Delta t \boldsymbol{f}(\boldsymbol{X}^n) \tag{20}$$

with initial conditions $J^0 = [0, I]^\top, \boldsymbol{X}^0 = \boldsymbol{Y}$. In our problem, we have

$$\boldsymbol{f}(\boldsymbol{X}) = \begin{bmatrix} \dfrac{\partial H}{\partial \boldsymbol{q}} \\ -\dfrac{\partial H}{\partial \boldsymbol{\alpha}} \end{bmatrix}, \ \ \boldsymbol{X} = \begin{bmatrix} \boldsymbol{q} \\ \boldsymbol{\alpha} \end{bmatrix}, \ \ \frac{\mathrm{d}\boldsymbol{f}(\boldsymbol{X}^n)}{\mathrm{d}\boldsymbol{X}^n} = \begin{bmatrix} \dfrac{\partial^2 H}{\partial \boldsymbol{q} \partial \boldsymbol{\alpha}} & \dfrac{\partial^2 H}{\partial \boldsymbol{\alpha}^2} \\ -\dfrac{\partial^2 H}{\partial \boldsymbol{q}^2} & -\dfrac{\partial^2 H}{\partial \boldsymbol{\alpha} \partial \boldsymbol{q}} \end{bmatrix} \ . \tag{21}$$

The entries of the Jacobian in (21) can be calculated explicitly. The analytic calculation of the Jacobian permits efficient solution of the nonlinear equation $\Phi(\boldsymbol{A}; \boldsymbol{P}) = \boldsymbol{Q}$ using the NAG nonlinear solver `nag_nlin_sys_sol` [9].

## 4 Comparison of Techniques

*Numerical Optimization Approach.* The inner iterations of the optimization method of Lancelot B use a minimization of a quadratic model function for which an approximate minimum is found. This leads to a model reduction by solving one or more quadratic minimization problems, requiring a solution of a sequence of linear systems. The Lancelot optimization package allows a choice of linear solver from 12 available options as described in [7]. The tests use 10

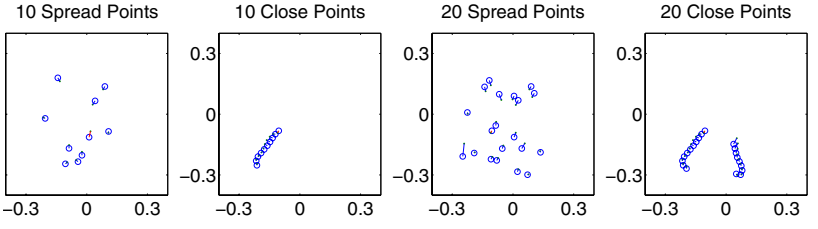



**Fig. 2.** Test cases for testing the linear solvers showing $\boldsymbol{P}_i$ as circles and $\boldsymbol{Q}_i$ as points

time steps on four selected test cases, as illustrated in Fig. 2, two comprising knotpoints taken from adjacent points in a data-set of 123 knotpoints hand-annotated on an MRI brain image, and two comprising points taken equally spaced throughout the data-set. We see the results of the experiment in Table 1. In general, the problems with the points closer together in the domain are hardest to solve. However, the method using the Munksgaard preconditioner, for instance, performs better with 10 close points than with 10 points spread through the domain. The expanding band preconditioner performs the best over the four test cases, but still shows unexpected behaviour in the 20 point tests. It is clear from these experiments that it is difficult to predict how an individual linear solver will perform on a particular problem. None of the linear solvers resulted in convergence for the case involving all of the 123 knotpoints.

In order to understand this behaviour better, we explore the change in condition number of the interpolation matrix with respect to the minimum separation between knotpoints. First, we examine the interpolation matrix for the clamped-plate spline. Computing the clamped-plate spline requires solving a linear system involving an interpolation matrix, G, constructed of biharmonic Green's functions in the same manner as that in which we constructed matrix $\mathcal{A}$ in Sec. 3, namely $\mathrm{G}_{ij} = G(\boldsymbol{q}_i, \boldsymbol{q}_j)$, where $G(\cdot, \cdot)$ denotes the $d$-dimensional biharmonic Green's function and $\boldsymbol{q}_i, \boldsymbol{q}_j$ are $d$-dimensional knotpoint position vectors.

**Table 1.** Time taken to converge (in seconds)

| Solver | 10 Spread | 10 Close | 20 Spread | 20 Close | Total |
|---|---|---|---|---|---|
| Conjugate Gradient | 5.78 | 4.86 | 54.04 | 71.07 | 135.74 |
| Diagonal Band | 24.38 | 46.40 | 366.98 | 785.47 | 1223.23 |
| Expanding Band | 0.89 | 1.69 | 10.00 | 6.35 | 18.93 |
| Munksgaard | 22.98 | 15.04 | 128.72 | 158.95 | 325.68 |
| Lin-More | 4.53 | 4.62 | 49.26 | 89.20 | 147.61 |
| Multifrontal | 1.42 | 2.09 | 76.07 | 21.00 | 100.58 |

This interpolation matrix is also a key feature of the Hessian of the augmented Lagrangian, as discussed in Sec. 2. In Fig. 3, we see the change in condition number for the interpolation matrix for two parallel knotpoint paths. From the literature, [10], we expect the condition number in the 2-norm, $\kappa_2(\mathrm{G})$ of the interpolation matrix to vary with the minimum separation of the knotpoints in a manner bounded above by $\min_{i,j} \|\boldsymbol{q}_i - \boldsymbol{q}_j\|^{-\beta}$, $\beta \in \mathbb{R}^+$. Accordingly, for comparison, we calculate and plot $\alpha \min_{i,j} \|\boldsymbol{q}_i - \boldsymbol{q}_j\|^{-\beta}$. We suspect the poor performance of the linear solvers in the test cases is due to very large condition numbers for small separations.

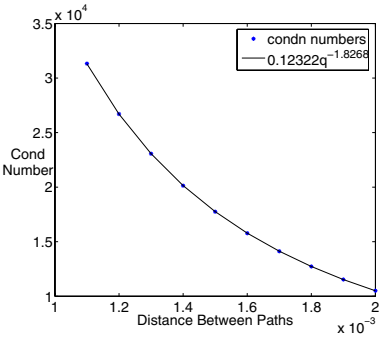



**Fig. 3.** Condition numbers for knotpoints with decreasing separation

*Classical Mechanics Approach.* Experiments show that the Hamiltonian implementation can solve all of the test cases illustrated in Fig. 2 in less than one second, whereas the Lancelot B implementation takes over 18 seconds to solve the test cases. We see the effect of an increase in the number of time steps in Fig. 4, comparing results using a numerical Jacobian with those using a true Jacobian. We see that the performance of the method using the true Jacobian is much superior to that using a numerical Jacobian, both in terms of function evaluations and of time. These tests used the first 60 knot points of the 123 point set. In Fig. 4, we see the effect on performance of the Hamiltonian implementation under an increase in the number of knotpoints, both with a numerical Jacobian and with a user-supplied analytic Jacobian, as described above. The knotpoints

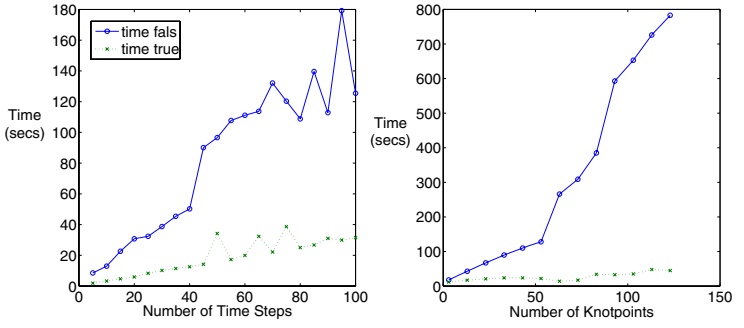



**Fig. 4.** Varying the number of time steps (l) and knotpoints (r) for the Hamiltonian

are taken as consecutive excerpts from the same set of 123 points as for the test sets above and 20 time steps are used. The speed of convergence of the method improves by a factor of approximately 4 when there is a user-supplied Jacobian. Notice that solving for the full 123 point set takes less than 40 seconds using this method.

## 5 Conclusions

The Galahad implementation was developed by the authors as an improvement to Marsland and Twining's original MATLAB method [1] for calculating the GIS and showed significant improvement over the MATLAB implementation. The Hamiltonian method shows an impressive improvement over both of these methods. We believe that the Galahad method shows disappointing performance due to the lack of a preconditioner suitable for the ill-conditioning of the interpolation matrix. The Hamiltonian method for computing the GIS dramatically outperforms the Lancelot B implementation over the test set of real data. It is clear that exact Jacobians should be supplied to the Hamiltonian implementation to give efficient performance. We see from the experiments carried out that, with exact Jacobians provided, the performance of the Hamiltonian method is superior to the performance of previous methods.

## References

1. Marsland, S., Twining, C.: Measuring geodesic distances on the space of bounded diffeomorphisms. In: British Machine Vision Conference (BMVC). (2002)
2. Modersitzki, J.: Numerical methods for image registration. Oxford University Press, New York (2004)
3. Camion, V., Younes, L.: Geodesic interpolating splines. In: EMMCVPR02. (2002) 513 ff.
4. Marsland, S., Twining, C.J.: Constructing diffeomorphic representations for the groupwise analysis of nonrigid registrations of medical images. IEEE Trans. Med. Imaging **23** (2004) 1006–1020

5. Cheney, W., Light, W.: A course in approximation theory. Brooks/Cole (1999)
6. Boggio, T.: Sulle funzioni di green d'ordine m. Circolo Matematico di Palermo **20** (1905) 97–135
7. Gould, N.I.M., Orban, D., Toint, P.L.: Galahad, a library of thread-safe fortran 90 packages for large-scale nonlinear optimization. ACM Trans. Math. Softw. **29** (2003) 353–372
8. Arnold, V.I.: Mathematical methods of classical mechanics. 2 edn. Volume 60 of Graduate Texts in Mathematics. Springer Verlag, New York (1989) 508 pages.
9. Numerical Algorithms Group http://www.nag.co.uk/numeric/FN/manual/: NAG Manual. (2004)
10. Schaback, R.: Error estimates and condition numbers for radial basis function interpolation. Advances in Computational Mathematics **3** (1995) 251–264

# Combining Registration and Abnormality Detection in Mammography

Mohamed Hachama, Agnès Desolneux, and Frédéric Richard

MAP5, University Paris 5,
45, rue des Saints Pères,
75006 Paris, France
{hachama, desolneux, richard}@math-info.univ-paris5.fr

**Abstract.** Usually, image registration and abnormality detection (e.g. lesions) in mammography are solved separately, although the solutions of these problems are strongly dependent. In this paper, we introduce a Bayesian approach to simultaneously register images and detect abnormalities. The key idea is to assume that pixels can be divided into two classes: normal tissue and abnormalities. We define the registration constraints as a mixture of two distributions which describe statistically image gray-level variations for both pixel classes. These mixture distributions are weighted by a map giving probabilities of abnormalities to be present at each pixel position. Using the Maximum A Posteriori, we estimate the deformation and the abnormality map at the same time. We show some experiments which illustrate the performance of this method in comparison to some previous techniques.

## 1 Introduction

Mammograms are often interpreted by comparing left and right breasts or successive mammograms of a same patient. Such comparisons help radiologists to locate suspicious differences which indicate the presence of some abnormalities [1]. Several Computer-Aided Diagnosis (CAD) systems have also used image comparisons for detecting abnormalities [2,3,4].

In these systems, a difference image is used to compare two images. This difference image is obtained by simple subtraction [2], weighted subtraction [3] or nonlinear subtraction [4]. Then it is thresholded to extract suspicious regions. However, the image comparison is not straightforward due to additional image dissimilarities which are related to sensor noise, different radiation exposure, and variation of breast positioning and compression and which cause high false-negative rates in abnormality detection schemes. Image registration is commonly carried out to compensate for these normal differences. Hence, the success of the detection task based on image difference depends on the preliminary registration process.

On the other hand, the registration problem is usually expressed as a minimization of an energy composed of a regularization term and a similarity term.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 178–185, 2006.

The definition of the similarity criterion relies on the nature of image gray-level dependencies [5]. For instance, the Sum of Square Differences (SSD) is often used whenever gray-level values are approximately the same in the two images to be registered. However, the presence of pathologies in mammograms such as lesions invalidates gray-level dependency assumptions. Hence, abnormalities may distort registration constraints and cause registration errors. Incorporating some knowledge about abnormalities can improve the registration.

In [6], F. Richard proposed a registration technique which down-weights the influence of abnormalities in the computation of registration constraints. The similarity criterion used is related to M-estimation criteria, also applied for optical flow computation [7]. But, the M-estimation approaches characterize abnormalities as pixels generating large image differences, which is not always the case. A more general approach consists of using mixture-based models, in which abnormalities are represented by a probability distribution, as it was done for the optical flow estimation by Jepson and Black [8], and for image registration by Hasler et al. [9], and by Hachama et al. [10].

In this paper, we present a mixture-based technique related to these previous works. The main feature of our model is the definition of a probability lesion map, which weights the mixture distributions at each pixel position by a probability to belong to a lesion. In this manner, we can interleave the registration and abnormality detection and thus take proper advantage of the dependence between the two processes.

The mixture-based technique and its mathematical formulation are presented in Section 2. In Section 3, we illustrate the method behavior on some examples and compare it with some classical techniques.

## 2 The Mixture-Based Technique

Let $I$ and $J$ be two images of the same size $(M, N)$, having gray-level values in the set $\{0, ..., 255\}$ and defined on a discrete grid $\Omega_d = \{(\frac{i}{M-1}, \frac{j}{N-1}), (i, j) \in \{0, ..., M-1\} \times \{0, ..., N-1\}\}$ associated to $\Omega = [0, 1]^2$. Image coordinates are matched using applications $\phi$ which map $\Omega_d$ into itself. Usually, registering the source image $I$ and the target image $J$ consists of finding an application $\phi$ which is such that the deformed image $I_\phi = I \circ \phi$ is "similar" to the target image $J$.

We assume that lesions may be present in the images. Let $L$ be the lesion map which associates to each pixel of $\Omega_d$ its probability to belong to a lesion. In the following, we formulate a bayesian model which allows us to estimate simultaneously the deformation $\phi$ and the lesion map $L$. Thus, we can solve the problems of image registration and abnormality detection at the same time.

### 2.1 Bayesian Formulation

Our formulation follows the Bayesian framework for image analysis laid out in [11]. Assuming that images, transformations and lesion maps are realizations of some random fields, Bayes rule can be expressed as

$$p(\phi, L|I, J) \;=\; \frac{p(I, J|\phi, L)\, p(\phi, L)}{p(I, J)}.$$

The Bayes rule allows us to write the posterior distribution $p(\phi, L|I, J)$ which contains the information about the unknowns $\phi$ and $L$, in terms of the prior $p(\phi, L)$ and the likelihood $p(I, J|\phi, L)$. The prior contains information about the most likely deformations and possible forms of lesions, namely, their morphology and spatial configuration. The relation between registered images is encapsulated in the likelihood term. The probability $p(I, J)$ is constant because it only depends on the observed fields $I$ and $J$.

As a simplification, we assume that the deformation $\phi$ and the lesion map $L$ are independent. In fact, lesions could generate specific local deformations that we choose to neglect. Thus, the Bayes rule can be written as

$$p(\phi, L|I, J) \;\alpha\; p(I, J|\phi, L)\, p(\phi)\, p(L).$$

We can estimate the pair $(\phi, L)$ as the solution of the Maximum A Posteriori:

$$(\widetilde{\phi}, \widetilde{L}) = \arg\max_{(\phi,L)} \; p(I, J|\phi, L)\; p(\phi)\; p(L)\,. \tag{1}$$

To ensure that the transformations remain smooth, we assume that they arise from the Gibbs distribution:

$$p(\phi) = \frac{1}{Z_1} e^{-H_d(\phi)}\;, \tag{2}$$

where $Z_1$ is a normalization constant, and $H_d$ is a discrete elasticity potential [12] (a continous version is given by Equation (9)). We also assume that the lesion map arises from a Gibbs distribution:

$$p(L) = \frac{1}{Z_2} e^{-R_d(L)}\;, \tag{3}$$

where $Z_2$ is a normalization constant, and $R_d$ is a discrete energy of regularization. We use in this paper an energy restricting the amount of abnormal pixels in the images via a real parameter $\alpha_L$ :

$$R_d(L) = \alpha_L \sum_{x \in \Omega_d} L(x)\,.$$

More specific terms can be defined to describe the spatial configurations of each type of lesion. We will investigate the use of such energies in the future.

In order to define the likelihood $p(I, J|\phi, L)$, we assume that, given the transformation $\phi$, the probability of the pair of images $(I, J)$ depends only on the registered images $(I_\phi, J)$ and that pixels are independent. Hence, we can write

$$p(I, J|\phi, L) = \prod_x p(I_\phi(x), J(x)|L(x)). \tag{4}$$

The probability of the pair $(I_\phi(x), J(x))$ depends on the class of the pixel $x$. Each class is characterized by a probability distribution, denoted by $p_{\mathrm{N}}$ for the normal tissue and $p_{\mathrm{L}}$ for the lesion. Thus, the probability distribution $p(I_\phi(x), J(x)|L(x))$ can be defined as a mixture of the two class distributions:

$$p(I_\phi(x), J(x)|L(x)) = (1 - L(x))p_{\mathrm{N}}(I_\phi(x), J(x)) + L(x)p_{\mathrm{L}}(I_\phi(x), J(x)). \quad (5)$$

The value of the lesion map $L$ at location $x$ is used to weight both class distributions. In what follows, we present the distributions $p_{\mathrm{N}}$ and $p_{\mathrm{L}}$ we used in experiments.

**The Normal Tissue Class.** Normally, gray-level values of registered images should be exactly the same at corresponding positions. But, in practice, these gray-level values usually differ because of noise or different image acquisition parameters. Assuming that these variations have a discrete Gaussian distribution with mean 0 and variance $\sigma^2$ ($\sigma = 15$ in the experiments), we can define $p_{\mathrm{N}}$ as

$$p_{\mathrm{N}}(I_\phi(x), J(x)) = \frac{1}{C_1} \exp(-\frac{|\, I_\phi(x) - J(x)\, |^2}{2\sigma^2}), \quad (6)$$

where $C_1$ is the normalization constant.

**The Lesion Class.** The definition of the lesion distribution is a difficult task. Each type of lesion requires the definition of a specific distribution. For the sake of simplicity, we assume that a lesion is present in the target image $J$. We characterize the lesion just as an area which is brighter in the target image than it is in the source image, defining the following distribution:

$$p_{\mathrm{L}}(I_\phi(x), J(x)) = \begin{cases} 0, & \text{if } I_\phi(x) > J(x) \\ \frac{1}{C_2}, & \text{otherwise,} \end{cases} \quad (7)$$

where $C_2$ is the normalization constant.

### 2.2 Numerical Resolution

Up to now, we have formulated a Bayesian registration model in a discrete setting. We now transform the discrete model into a continuous model so as to be able to use variational resolution techniques. First, we rewrite the MAP estimate (Equation (1)) as the minimization of the negative-log function

$$E_d(\phi, L) = -\log(p(\phi)) - \log(p(L)) - \log(p(I, J|\phi, L)).$$

Then, using Equation (4) and Gibbs distributions (2) and (3), we get

$$E_d(\phi, L) = H_d(\phi) + R_d(L) - \sum_{x \in \Omega_d} \log(p(I_\phi(x), J(x)|L(x))) + K,$$

where $K$ is a constant. Next, following approaches in [13,12], we define a continuous expression of this energy, by interpolating all functions by the finite element method and replacing sums on the pixel grid $\Omega_d$ by integrals on $\Omega$:

$$E(\phi, L) = H(\phi) + R(L) - \int_{\Omega} \log(p(I_{\phi}(x), J(x))) \, dx, \tag{8}$$

where the probability distribution $p(I_{\phi}(x), J(x))$ is the obtained continuous version of the mixture distribution given by Equation (5). $H(\phi)$ is the elasticity potential defined as

$$\sum_{i,j=1,2} \int_{\Omega} [\lambda \frac{\partial u_i(x)}{\partial x_i} \frac{\partial u_j(x)}{\partial x_j} + \mu (\frac{\partial u_i(x)}{\partial x_j} + \frac{\partial u_j(x)}{\partial x_i})^2] dx, \tag{9}$$

where $u = \phi - id$, and $\lambda$ and $\mu$ are the Lame elasticity constants. The term $R(L)$ is the following energy:

$$R(L) = \alpha_L \int_{\Omega} L(x) \, dx \, .$$

As in [14,6], we use a gradient descent algorithm on the energy $E$ and finite elements to approximate solutions of the minimization problem. We use a variable change $L = \frac{1}{1+e^{-M}}$ so as to satisfy the constraints $0 \leq L(x) \leq 1$ and to be able to differentiate the energy with respect to the second variable $M$.

## 3 Results

In this section, we illustrate the characteristics of the mixture model by comparing its performance to those of the SSD technique [12], and the M-estimator based technique proposed in [6]. We applied algorithms to a pair of bilateral mammograms (case 21 of the MIAS database [15]), for which the target image contains a lesion (bright circular region at the bottom of Image (1-b)).

*Registration results.* Registrations obtained with the SSD and M-estimation techniques tend to incorrectly match the lesion and the bright tissue in the source image and thus reduce image differences due to the lesion (Images (1-d) and (1-e)). This is corrected by the mixture-based technique which registers the images correctly while preserving differences due to the lesion (Image (1-f)).

*Detection results.* We compare lesion binary images obtained with the three techniques. For the SSD and the M-estimation techniques, lesion binary images are obtained by thresholding the image difference generated by the adaptively weighted subtraction [3]. The fact that abnormal pixels tend to have relatively higher gray-level values is used to weight the difference between a pair of pixels by gray-level value of the pixel of the image $J$. For the mixture-based method, we set $\alpha_L = 0.1$ and threshold the lesion map. The thresholds are chosen so as to have the same amount of abnormal pixels in the three lesion binary images obtained. Figure 2 shows the lesion binary images obtained with the three techniques for different amounts of abnormal pixels.

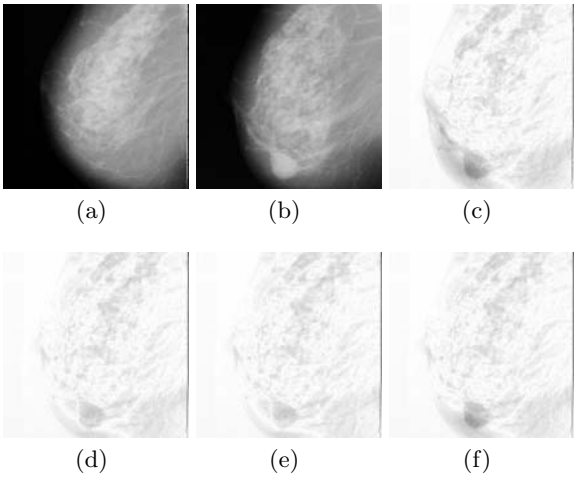



**Fig. 1.** Registration of bilateral mammograms. (a) Source image I, (b) Target image J, (c) The difference between the images before registration. The difference between the images after the registration using (d) the SSD method, (e) the M-estimation method, (f) the mixture-based method.

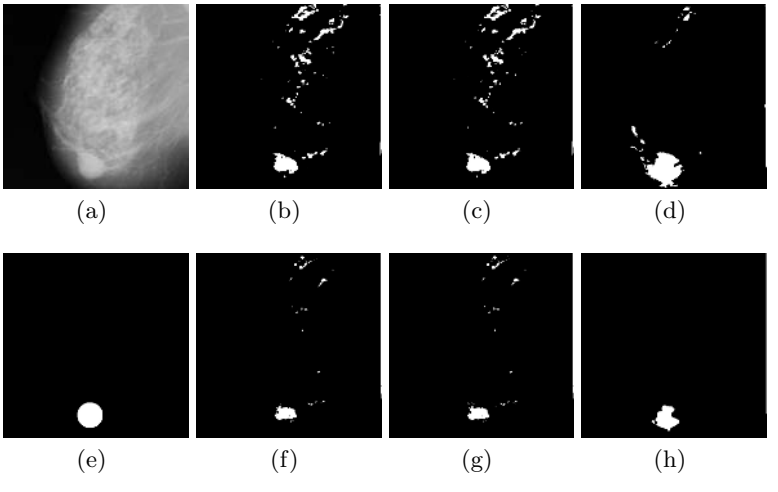



**Fig. 2.** The detection results. (a) The target image containing the lesion. For 10550 abnormal pixels, the results obtained with the (b) SSD method, (c) M-estimation based method, (d) The mixture-based method. (e) The expert segmented lesion. For 4180 pixels, the results obtained with the (f) SSD method, (g) M-estimation based method, (h) The mixture based method.

For evaluating and comparing the three algorithms without the influence of a threshold value, we have represented on Figure 3 the FROC curves obtained with the three methods. The FROC curve plots the sensitivity (fraction of detected true positives calculated by using the expert segmented image) as a function of the number of false positives. For the mixture-based technique, we have obtained similar FROC curves for different values of the weight $\alpha_L$. We have represented the FROC curve obtained when $\alpha_L = 0.1$.

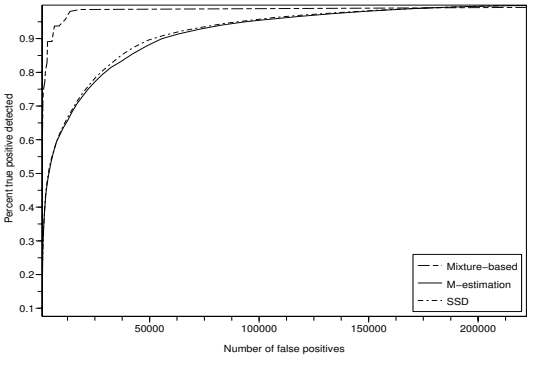



**Fig. 3.** FROC Curves for the three detection methods

As observed on Figure (3), the FROC curve associated to the mixture-based method is the highest. So, the detection by the mixture-based technique is more sensitive. For instance, for 10000 false positive pixels (2% of image pixels), the detection rate grows from 0.632 for the SSD and 0.627 for the M-estimation based method, to 0.947 for the mixture based method.

## 4 Conclusion

We have presented a method for joint mammogram registration and abnormality detection. Thanks to this combined approach, the mixture-based method improves the mammogram registration and increase the sensitivity of lesion detection. In the future, we will focus on how to design a lesion model for different types of lesions, and on the estimation of the distribution parameters for both lesion and normal tissue classes. Furthermore, we plan to apply the mixture method to a full mammogram database.

**Acknowledgments.** This work was supported by the grant "ACI Young Researchers 2003 (French Ministry of Research)" No. 9060.

## References


1. L. Tabar and P. Dean, *Teaching atlas of mammography*, Thieme Inc., Stuttgart, 1985.
2. M. Sallam, G. Hubiak, K. Bowyer, and L. Clarke, "Screening mammogram images for abnormalities developing over time," in *IEEE Nuclear Science Symposium and Medical Image Conference*, 1992, pp. 1270–1272.
3. M.Y. Sallam and K. Bowyer, "Registration and difference analysis of corresponding mammogram images," *Medical Image Analysis*, vol. 3, no. 2, pp. 103–118, 1999.
4. F-F. Yin, M.L. Giger, K. Doi, C.E. Metz, C.J. Vyborny, and R.A. Schmidt, "Computerized detection of masses in digital mammograms: Analysis of bilateral-subtraction images," *Medical Physics*, vol. 18, pp. 995–963, 1991.
5. A. Roche, G. Malandain, and N. Ayache, "Unifying Maximum Likelihood Approaches in Medical Image Registration," *International Journal of Computer Vision of Imaging Systems and Technology*, vol. 11, pp. 71–80, 2000.
6. F. Richard, "A new approach for the registration of images with inconsistent differences," in *Proc. of the Int. Conf. on Pattern Recognition, ICPR*, Cambridge, UK, 2004, vol. 4, pp. 649–652.
7. N. Sebe and M.S. Lew, *Robust Computer Vision: Theory and Applications*, vol. 26, Series: Computational Imaging and Vision, 2003.
8. A. Jepson and M.J. Black, "Mixture models for optical flow computation," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 1993, pp. 760–761.
9. D. Hasler, L. Sbaiz, S. Susstrunk, and M. Vetterli, "Outlier modeling in image matching," *IEEE Trans. on Patt. Anal. and Match. Intell.*, pp. 301–315, 2003.
10. M. Hachama, F. Richard, and A. Desolneux, "A mammogram registration technique dealing with outliers," *University Paris 5, MAP5 Technical Report, www.math-info.univ-paris5.fr/map5/publis/*, January 2006.
11. D. Mumford, "The bayesian rationale for energy functionals," *Geometry-driven Diffusion in Comp. Vis., Kluwer Academic Publ*, vol. 46, pp. 141–153, 1994.
12. F. Richard, "A comparative study of markovian and variational image-matching techniques in application to mammograms," *Pattern Recognition Letters*, vol. 26, no. 12, pp. 1819–1829, 2005.
13. M.I. Miller, G. Christensen, Y. Amit, and U. Grenander, "Mathematical textbook of deformable neuroanatomies," in *Proc. Natl. Acad. Sc.*, USA, 1993, pp. 11944–11948.
14. F. Richard and L. Cohen, "Non-rigid image registration with free boundary constraintes: application to mammography," *Journal of Computer Vision and Image Understanding*, vol. 89(2), pp. 166–196, 2003.
15. J. Suckling, J. Parker, and D. Dance, "The MIAS digital mammogram database," *In Proc. of the 2nd Int. Workshop on Digital Mammography, England*, july 1994.

# Point Similarity Measures Based on MRF Modeling of Difference Images for Spline-Based 2D-3D Rigid Registration of X-Ray Fluoroscopy to CT Images

Guoyan Zheng[1], Xuan Zhang[1], Slavica Jonić[2, 3], Philippe Thévenaz[2], Michael Unser[2], and Lutz-Peter Nolte[1]

[1] MEM Research Center, University of Bern, Stauffacherstrasse 78, CH-3014, Bern, Switzerland
Guoyan.Zheng@MEMcenter.unibe.ch
[2] Biomedical Imaging Group, École polytechnique fédérale de Lausanne (EPFL), CH-1015, Lausanne VD, Switzerland
[3] Institut de Minéralogie et de Physique des Milieux Condensés, Université Pierre et Marie Curie, F-75015 Paris, France

**Abstract.** One of the main factors that affect the accuracy of intensity-based registration of two-dimensional (2D) X-ray fluoroscopy to three-dimensional (3D) CT data is the similarity measure, which is a criterion function that is used in the registration procedure for measuring the quality of image match. This paper presents a unifying framework for rationally deriving point similarity measures based on Markov random field (MRF) modeling of difference images which are obtained by comparing the reference fluoroscopic images with their associated digitally reconstructed radiographs (DRR's). The optimal solution is defined as the maximum a posterior (MAP) estimate of the MRF. Three novel point similarity measures derived from this framework are presented. They are evaluated using a phantom and a human cadaveric specimen. Combining any one of the newly proposed similarity measures with a previously introduced spline-based registration scheme, we develop a fast and accurate registration algorithm. We report their capture ranges, converging speeds, and registration accuracies.

## 1 Introduction

One of the main factors that affect the accuracy of intensity-based 2D-3D registration is the similarity measure, which is a criterion function that is used in the registration procedure for measuring the quality of image match. An extensive study of six similarity measures applied specifically to 2D-3D registration has been performed by Penney *et al*. [1]. The similarity measures considered by the authors were: normalized cross-correlation [2], entropy of the difference image [3], pattern intensity [4], mutual information [5], gradient correlation [6], and gradient difference [1]. Using the fiducial markers to get the "gold-standard" registration, the authors ranked these measures based on their accuracy and robustness. They found that pattern intensity was one of the two similarity measures that were able to register accurately and robustly, even when soft tissues and interventional instruments were present in the X-ray images. Unfortunately, pattern intensity was designed by using some heuristic rules [4].

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 186 – 194, 2006.

This work formulates a MRF model on the difference images obtained by comparing the input fluoroscopic images with their associated DRR's. The optimal solution is defined as the MAP estimated of the MRF. By using this unifying MAP-MRF framework, we can derive new point similarity measure in a rational way. The optimization of each individual similarity measure derived from this framework leads to optimal registration. We point out that two previously published similarity measures, i.e., sum-of-squared-difference (SSD) [7] and pattern intensity [4], can be also derived from this framework.

The remainder of this paper is organized as follows. Section 2 briefly introduces the 2D-3D registration scheme used in this paper. Section 3 describes the derivation of point similarity measures based on MRF modeling of the difference images. Section 4 presents the experimental results, followed by conclusions in Section 5.

## 2 Spline-Based 2D-3D Registration Scheme

The 2D-3D registration scheme used in this paper is based on a recently introduced spline-based multi-resolution 2D-3D registration scheme [7, 8]. This scheme follows the computation framework of intensity-based methods. Given a set of X-ray images and a CT volume, it iteratively optimizes the six rigid-body parameters describing the orientation and the translation of the patient pose, by generating and comparing floating DRR's with the reference X-ray images using appropriate similarity measure. The differences between this method and other intensity-based methods lie in [7]: 1) a cubic-splines data model was used to compute the multi-resolution data pyramids for both CT volume and X-ray images, the DRR's, as well as the gradient and the Hessian of the cost function; 2) a Marquardt-Levenberg non-linear least-squares optimizer was adapted to a multi-resolution context. The registration was performed from the coarsest resolution until the finest one. The accuracy of this method depends on the chosen similarity measure. Previously, accuracy of approximately $1.4 \pm 0.2$ mm when SSD was used [7] has been reported.

## 3 Deriving Point Similarity Measures Based on MRF Modeling of Different Images

To find an optimal registration transformation we cast the problem into a Bayesian framework of MAP-MRF estimate. We thus follow the four steps of the MAP-MRF estimate [9].

1. Construction of a prior probability distribution $p(T)$ for the registration transformation T matching the reference X-ray images to the floating DRR's.
2. Formulation of an observation model $p(D \mid T)$ that describes the distribution of the observed difference images D by comparing the reference X-ray images and the floating DRR's given any particular realization of the prior distribution.

3. Combination of the prior and the observation model into the posterior distribution by Bayes theorem

$$p(T \mid D) \propto p(D \mid T)p(T) \tag{1}$$

4. Drawing inference based on the posterior distribution.

### 3.1 Prior Distribution

One advantage of formulating 2D-3D registration according to Bayesian framework is that we are able to specify a prior distribution for each configuration of registration parameter space. In this paper, we don't take advantage of this property. We treat all parameter configurations equally. Due to the Euler angle based parameterization of rotation in our approach, $p(T)$ are a uniform distribution. But it is possible to use this property to favor certain transformations when different parameterization forms such as quaternion are used.

### 3.2 Observation Model

Given a realization of the prior distribution, the observation model *p(D|T)* describes the conditional distribution of the observed difference images *D*. By specifying an observation model we may favor a transformation that establishes matching between regions of similar properties. By modeling the difference image *D* as a MRF with respect to the *r*th order neighborhood system $N = \{N_{i,j}^r\}$ we can derive the energy function for the observation model as:

$$E(D \mid T) = \sum_{q=1}^{Q} [\alpha \sum_{i,j}^{I,J} V(d_{i,j}) + (1-\alpha) \sum_{i,j}^{I,J} \frac{1}{card(N_{i,j}^r)} \sum_{(i',j') \in N_{i,j}^r} V(d_{i,j}, d_{i',j'})] \tag{2}$$

where $Q$ is the number of images and $I \times J$ is the size of each image. The first term is the potential function for single-pixel cliques and the second term is the potential function for all other cliques. $\alpha \in [0:1]$ weights the influence of these two terms. $card(N_{i,j}^r)$ means to compute the number of pixels in neighborhood $N_{i,j}^r$.

The selection of the potential functions in Eq. (2) is a critical issue in MRF modeling [9]. As pointed out below, its selection decides the form of similarity measure.

The computation of the difference images also plays an important role in the present framework. In [4], an adjustable scaling parameter was used to build the difference images. To eliminate this parameter, Jonić *et al.* [7] tries to normalize the intensity range of the input reference fluoroscopic images and that of the corresponding DRR's by removing their mean and then dividing by their standard deviation. In this paper, we use a similar method. But unlike in [7], where the mean and the standard deviation were computed from the complete region of interest (ROI), we compute them only using those pixels in the neighborhood $N_{i,j}^r$.

### 3.3 MAP Estimate

The posterior conditional probability distribution is given by:

$$p(T \mid D) \propto \exp(-E(D \mid T)) \tag{3}$$

In search for the MAP estimate:

$$\widehat{T} = \arg\max_T \; p(T \mid D) \tag{4}$$

To illustrate how to derive similarity measures using the present framework, two examples of previously published similarity measures are given as follows.

**Sum-of-Squared-Difference (SSD):** It can be derived from Eq. (2) by specifying $\alpha = 1$ and $V(d_{i,j}) = d_{i,j}^2$.

**Pattern Intensity:** the pattern intensity proposed in [4] is written in the form:

$$P_{r,\sigma} = \sum_{i,j} \frac{1}{card(N_{i,j}^r)} \sum_{(i',j') \in N_{i,j}^r} \frac{\sigma^2}{\sigma^2 + (d_{i',j'} - d_{i,j})^2} \tag{5}$$

where $r$ and $\sigma$ are two parameters to be experimentally determined. $N_{i,j}^r$ is a neighborhood with radius $r$. It can be derived from the present framework by specifying $\alpha = 0$ and using following pairwise clique potential function:

$$V(d_{i,j}, d_{i',j'}) = -\frac{1}{1 + \dfrac{(d_{i',j'} - d_{i,j})^2}{\sigma^2}} \tag{6}$$

where $d_{i',j'}$ is a pixel in the neighborhood $N_{i,j}^r$.

### 3.4 Deriving New Point Similarity Measures

More generally, by choosing different neighborhood system and by specifying different clique potential functions that incorporates different *a priori* constraints, we can derive different new similarity measures.

**Isotropic rth order neighborhood system and pairwise potential function with 1st order smoothness constraint (INrS1):** It is defined using following equation:

$$\alpha \sum_{i,j}^{I,J} d_{i,j}^2 + (1-\alpha) \sum_{i,j}^{I,J} \frac{1}{card(N_{i,j}^r)} \sum_{(i',j') \in N_{i,j}^r} (d_{i',j'} - d_{i,j})^2 \tag{7}$$

It is actually a combination of **SSD** and a modified form of pattern intensity [10]. Following the suggestion in [4], we also choose r=3 pixels. From now on, we call this similarity measure **IN3S1**.

Two anisotropic similarity measures can be derived using following equation:

$$P_\alpha = \alpha \sum_{i,j}^{I,J} d_{i,j}^2 + (1-\alpha) \sum_{i,j}^{I,J} (d_{(i,j),x}^2 + d_{(i,j),y}^2) \tag{8}$$

where $d_{(i,j),x} = \frac{\partial d}{\partial x}(i,j)$; $d_{(i,j),y} = \frac{\partial d}{\partial y}(i,j)$ is the first derivatives of the difference image *D* along X and Y directions, respectively.

**Anisotropic 4-neighborhood system and potential functions with first order smoothness constraint (AN4S1)**: It computes the first derivative in Eq. (8) using 4-neightborhood system with following convolution masks:

- $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$ for the determination of $d_{(i,j),x}$ and
- $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^T$ for the determination of $d_{(i,j),y}$

**Anisotropic 8-neighborhood system and potential functions with first order smoothness constraint (AN8S1)**: It also computes the first derivative in Eq. (8) using 4-neightborhood system but with following convolution masks:

- $\begin{bmatrix} -\frac{1}{3} & 0 & \frac{1}{3} \\ -\frac{1}{3} & 0 & \frac{1}{3} \\ -\frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix}$ for the determination of $d_{(i,j),x}$ and
- $\begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$ for the determination of $d_{(i,j),y}$

## 4 Experiments

A phantom and a human cadaveric spine specimen together with their ground truths were used in our experiments. Both phantom and cadaveric specimen were scanned by a GE LightSpeed Ultra CT scanner (GE Healthcare, Chalfont St. Giles, United Kingdom) with same intra-slice solution (0.36 mm x 0.36 mm) but with different inter-slice thickness, 1.25 mm for the phantom and 2.5 mm for the cadaveric specimen, which resulted in volume dataset of size 512x512x93 volxels for phantom and 512x512x72 for the cadaveric specimen, respectively. The 2D projection images of both phantom and cadaveric specimen were acquired from a Siemens ISO-C C-arm (Siemens AG, Erlangen, Germany). They are calibrated and undistorted with custom-made software with high accuracy. The phantom was custom-made to simulate a good

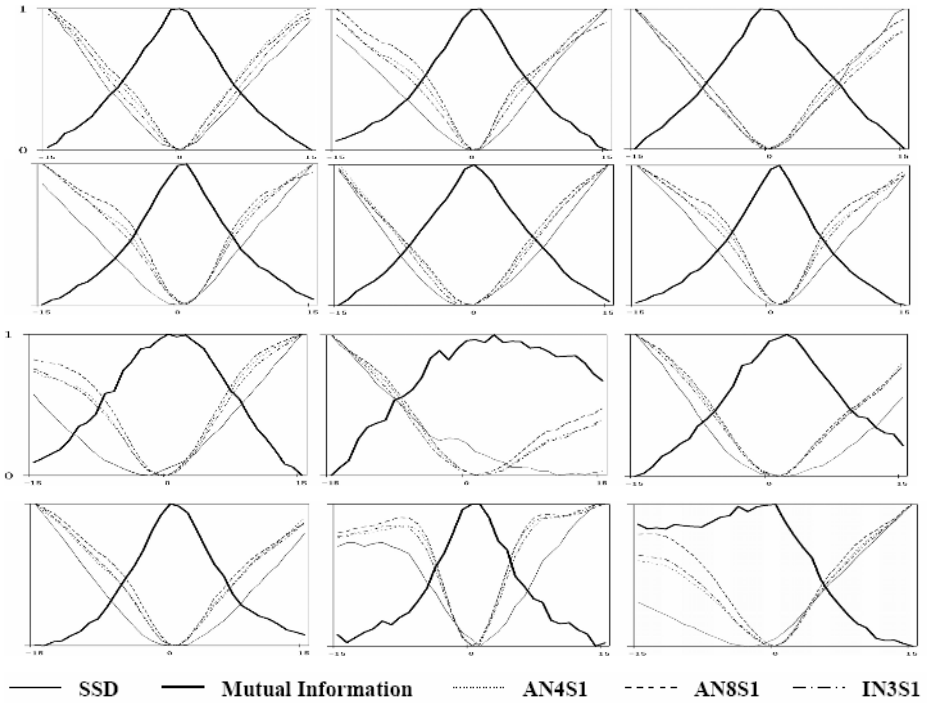



**Fig. 1.** Behavior of different similarity measures. Cut through the minimum of different similarity measures on the phantom data (the 1st and 2nd rows) as well as on the cadaveric spine specimen (the 3rd and 4th rows). The ordinate shows the value of different similarity measures (they are normalized to the range [0.0, 1.0]), which are given as functions of each rigid transformation parameter in the range of [-15°, 15°] or [-15 mm, 15 mm] away from the its ground truth ((1) 1st column of the 1st and 3rd rows: X rotation; (2) 2nd column of the 1st and 3rd rows: Y rotation; (3) 3rd column of the 1st and 3rd rows: Z rotation; (4) 1st column of the 2nd and 4th rows: X translation; (5) 2nd column of the 2nd and 4th rows: Y translation; (6) 3rd column of the 2nd and 4th rows: Z translation). Zero in each abscissa means the ground truth for that individual parameter, obtained by paired point matching based on fiducial markers.

condition. In contrast, projections of interventional instruments were present in the X-ray images of the cadaveric specimen to simulate a practical situation in image-guided therapy.

The ground truths were obtained by implanting fiducial markers. Both phantom and cadaveric specimen were equipped with infrared light emitting diodes (LEDs) markers to establish a patient coordinate system (P-COS) and was tracked using an optoelectronic position sensor (OptoTrak 3020, Northern Digital Inc., Waterloo, Canada). The actual locations of fiducial markers were digitized in P-COS using an optoelectronically tracked pointer and were matched to the corresponding points in CT volume dataset. The ground truths were then obtained using singular value decomposition with an accuracy of 0.52 mm for phantom and 0.65 mm for cadaver, respectively.

For all three newly derived similarity measures, the parameter $\alpha$ was chosen as 0.5. Each time, two nearly orthogonal C-arm images from the corresponding dataset were used for the experiments described below.

The first experiment was designed to compare the behaviors of the newly derived similarity measures to those of the published similarity measures such as **SSD** and mutual information. Though mutual information was ranked as least accurate in [1], other group [11, 12] later found that it performed reasonably well. The results were given in Figure 1. It was found that all similarity measures had similar behavior when tested on the phantom data but different behavior when tested on the cadaveric data. Those similarity measures derived from the present MAP-MRF framework showed a superior behavior compared to other two well-known similarity measures. More specially, the curves for the newly derived similarity measures have clear minima and are smoother, which is an important property to take the advantage of our 2D-3D registration scheme, which uses a gradient-based optimization technique. It is also evident that the behavior of mutual information is better than that of **SSD**.

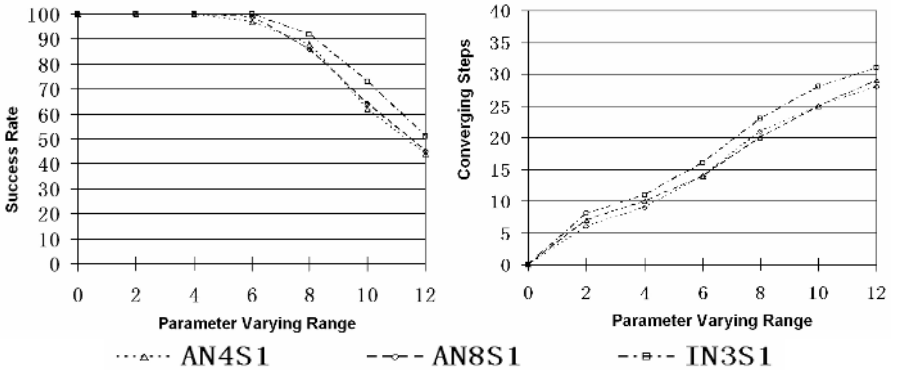



**Fig. 2.** Experimental results of capture ranges (left) and converging steps (right)

Combining any one of the similarity measures with the 2D-3D registration scheme described in Section 2, we developed a 2D-3D registration algorithm. The second experiment was designed to evaluate their capture ranges, converging steps, and registration accuracies of these registration algorithms. Based on the investigation results obtained in the first experiment, we only performed this experiment on the human cadaveric specimen dataset to compare the three newly derived similarity measures. For this purpose, we perturbed the ground truth transformation by randomly varying each registration parameter in the range of [-2°, 2°] or [-2mm, 2mm] to get 100 positions, and then another 100 positions in the range of [-4°, 4°] or [-4mm, 4mm], and so on until the final range of [-12°, 12°] or [-12mm, 12mm]. We then performed our registrations and counted how many times they converged for each range (when the target registration error (TRE) measured on those fiducial markers was less than 1.5 mm). The capture range was defined when there was at least 95% successful

**Table 1.** Results of registration accuracies

| Ranges | State | Target Registration Error (TRE) (mm) | | |
|---|---|---|---|---|
| | | **AN4S1** | **AN8S1** | **IN3S1** |
| [-2°, 2°] and [-2mm, 2mm] | Initial | 2.3 ± 0.7 | 2.3 ± 0.7 | 2.3 ± 0.7 |
| | Success Registration | 0.7 ± 0.2 | 0.7 ± 0.1 | 0.8 ± 0.1 |
| [-4°, 4°] and [-4mm, 4mm] | Initial | 4.6 ± 1.4 | 4.6 ± 1.4 | 4.6 ± 1.4 |
| | Success Registration | 0.7 ± 0.2 | 0.8 ± 0.1 | 0.8 ± 0.1 |
| [-6°, 6°] and [-6mm, 6mm] | Initial | 6.9 ± 2.2 | 6.9 ± 2.2 | 6.9 ± 2.1 |
| | Success Registration | 0.7 ± 0.1 | 0.8 ± 0.1 | 0.8 ± 0.1 |
| [-8°, 8°] and [-8mm, 8mm] | Initial | 8.9 ± 2.8 | 8.9 ± 2.8 | 8.9 ± 2.8 |
| | Success Registration | 0.7 ± 0.2 | 0.8 ± 0.1 | 0.8 ± 0.1 |
| [-10°, 10°] and [-10mm, 10mm] | Initial | 10.4 ± 3.6 | 10.4 ± 3.6 | 10.4± 3.6 |
| | Success Registration | 0.7 ± 0.2 | 0.8 ± 0.1 | 0.9 ±0 .2 |
| [-12°, 12°] and [-12mm, 12mm] | Initial | 11.6 ± 3.9 | 11.6 ± 3.9 | 11.6 ± 3.9 |
| | Success Registration | 0.7 ± 0.2 | 0.8 ± 0.1 | 0.9 ± 0.2 |

rate. The experimental results on capture ranges and converging steps are given in Figure 2. The results on registration accuracies are shown in Table 1. It was found that **IN3S1** had larger capture range than other two similarity measures but it was also less accurate and required more steps to be converged.

## 5 Conclusions

In this paper, we introduced a unifying MAP-MRF framework to derive novel point similarity measures for 2D-3D registration of X-ray fluoroscopy to CT images. The derived novel point similarity measures had been evaluated using phantom and cadaver and the results showed that they provided satisfactory 2D-3D registration accuracy, even when interventional instruments were present.

## References

1. Penney G.P., Weese J., Little J.A., Desmedt P., Hill D.L.G., and Hawkes D.J.: A comparison of similarity measures for use in 2D-3D medical image registration. IEEE T Med Imaging, Vol. 17, No. 4, (1998) 586-595
2. Lemieux L., Jagoe R., Fish D.R., Kitchen N.D., and Thomas D.G.T.: A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. Med Phys, Vol. 21, No. 11, (1994) 1749-1760
3. Buzug T.M., Weese J., Fassnacht C., and Lorenz C.: Image registration: convex weighting functions for histogram-based similarity measures. Lecture Notes in Computer Science, Vol. 1205, Springer-Verlag, Berlin Heidelberg New York (1997) 203-212

4. Weese J., Buzug T.M., Lorenz C., Fassnacht C.: An approach to 2D/3D registration of a vertebra in 2D x-ray fluoroscopies with 3D CT images. Lecture Notes in Computer Science, Vol. 1205, Springer-Verlag, Berlin Heidelberg New York (1997) 119-128
5. Maes F., Collignon A., Vandermeulen D., Marchal G., and Suetens P. : Multi-modality image registration by maximization of mutual information. IEEE T Med Imaging, Vol. 16, No. 2, (1997) 187-198
6. Brown L.M.G. and Boult T.E.: Registration of planar film radiographs with computed tomography. IEEE Proceedings of MMBIA (1996) 42-51
7. Jonić S., Thévenaz P., Zheng G., Nolte L.-P., and Unser M.: An Optimized Spline-based registration of a 3D CT to a set of C-arm images. International Journal of Biomedical Imaging (In Press), 2006.
8. Jonić S., Thévenaz P., and Unser M.: Multiresolution-based registration of a volume to a set of its projection. Proceedings of the SPIE International Symposium on Medical Imaging: Image Processing (MI' 03), San Diego CA, USA, (2003), Vol. 5032, Part II, pp. 1049-1052
9. Li S.Z.: Markov random field modeling in computer vision. Springer-Verlag, Berlin Heidelberg New York, (1995).
10. Zheng G., Zhang X., and Nolte L.-P.: Assessing spline-based multi-resolution 2D-3D image registration for practical use in surgical guidance. Lecture Notes in Computer Science, Vol. 3150, (2004) 294-301
11. Zöllei L., Grimson E., Norbash A., and Wells W.: 2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators. Proceedings of CVPR'01, Volume 2, (2001) 679-703.
12. Russakoff D.B., Rohlfind T., and Maurer C.R. Jr.: Fast intensity-based 2D-3D image registration of clinical data using lighting fields. Proceedings of ICCV'03, (2003) 416-422.

# Clinical Application of a Semiautomatic 3D Fusion Tool Where Automatic Fusion Techniques Are Difficult to Use

M.E. Noz[1], G.Q. Maguire Jr.[2], M.P. Zeleznik[3], L. Olivecrona[4], H. Olivecrona[5], L. Axel[1], M.B. Srichai[1], L. Moy[1], and A. Murphy-Walcott[1]

[1] Dept. of Radiology, New York University School of Medicine, New York, NY 10016, USA
[2] School of Information and Communication Technology, Royal Institute of Technology, SE-164 40 Kista, Sweden
[3] RAHD Oncology Products, St. Louis, Mo. 63135, USA
[4] Dept. of Radiology, [5]Dept. of Orthopedics, Karolinska Hospital, SE-171 76 Stockholm Sweden

**Abstract.** The purpose of this paper is to demonstrate the clinical advantages of using semiautomatic volume registration where automatic registration is problematic due to large deformations, small bone anatomy, or extraneous structures. Examples are drawn from clinical cases of MRI/PET breast studies, CT angiography/SPECT cardiac studies, and total wrist arthroplasty. These types of studies should be contrasted with those involving the head, thorax, and pelvis where there is much less deformation and the existence of (some) large bones facilitates automatic matching.

## 1 Introduction

Volumetric registration (fusion) of certain types of 3D data sets can be potentially clinically useful, but difficult to achieve with automatic fusion tools. Automated tools can be employed where each volume contains predictable information, such as with magnetic resonance imaging (MRI)/computed tomography (CT) head scans without skull or soft tissue deformation [1,2], and where expected anatomical structures and the body shape are similar, although the registration process can take a considerable amount of time [3-5] and is generally limited to an affine transformation. However, automated registration is more difficult in other situations. One example is the fusion of low resolution radionuclide functional scans that lack good anatomical features with higher resolution anatomical data, because the tools can't always detect the data correlation. Another example is the fusion of CT and positron emission tomography (PET) scans [6], where these tools don't know a priori what *not* to consider; moreover the voxels themselves do not provide this information. The registration method used in this study allows the user to employ a priori knowledge to explicitly indicate structural correlation, and to implicitly ignore other features, such as a prosthesis, which exist in the volumes. It also allows for 3D deformation. The user chooses co-homologous points that are used to estimate a transformation between the two volumes, which brings these points into alignment. The purpose of this report is to

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 195 – 205, 2006.

demonstrate the clinical advantages of using semiautomatic volume registration [7] in three cases where automatic registration is difficult; in (1) breast imaging and (2) heart imaging, because of large deformations and lack of bony landmarks, and (3) wrist imaging, because of the need to preserve the spatial relationship of multiple small structures.

1. In many instances, MRI scans of the breast showing tissue differences have low specificity resulting in the need for further invasive procedures or a series of six month follow-up scans [8]. Because metabolism is generally increased in tumors, 18-F fluoro-deoxy-glucose (18-F FDG) PET scans can provide useful information about the local metabolism of breast tissue. We fused volumetric data sets of the breast from MRI scans with 18-F FDG PET scans to assess whether the additional information from the PET study could improve the diagnostic outcome. These scans lack anatomic landmarks or other correlated data required for automatic fusion.

2. Coronary artery disease (CAD) is a leading cause of morbidity and mortality. Non-invasive evaluation for CAD has traditionally been based on the assessment of myocardial perfusion (single photon emission computed tomography-SPECT/PET) and/or function (Echocardiography/MRI) during rest and stress conditions. Definitive diagnosis is based on invasive coronary angiography demonstrating obstructive lesions involving the coronary arteries. However, the hemodynamic significance of coronary artery stenosis is not always apparent on routine coronary angiography, and an assessment of the myocardial territory supplied by the coronary artery is often needed. With the advent of multi-detector computed tomography (MDCT) scanners coupled with ECG gating, non-invasive evaluation of the coronary arteries has become possible. Although MDCT coupled with the dynamics of contrast agent enhancement is a promising method for myocardial perfusion assessment, the combination of relatively high doses of radiation and the use of potentially harmful contrast agents limits its clinical applicability [9]. To combine the advantages of both anatomic coronary artery evaluation with myocardial perfusion assessment, volume data sets from resting computed tomography angiography (CTA) and rest/stress SPECT studies were fused and correlated for clinical applicability. The automatic fusion of the these data sets is hampered because the areas of high activity in the CTA scan (the coronary arteries) correspond to areas of low activity on the SPECT scan which enhances the myocardium.

3. The most common long-term failure mode of total wrist arthroplasty (TWA) is loosening of the distal prosthetic component [10]. If detection of loosening and subsequent wear is delayed, the resulting bone destruction and soft tissue reactions can become more advanced. This can make revision procedures more extensive, painful for the patient, and costly. Standard planar radiographs are commonly used today to evaluate loosening of TWA. However, the accuracy of measurements on radiographs is low. It has previously been shown that position and migration of the acetabular component in total hip arthroplasty (THA) can be detected with the use of CT image registration [11]. Since the wrist joint is more complicated than the hip, there are new problems encountered when fusing TWA volumes, the most prominent of which is the small bone-to-prosthesis ratio (i.e., small bones, large metallic prosthesis). Again, auto-fusion techniques would be difficult here as the most prominent feature in the scans is the prosthesis. Unfortunately, it is precisely the possible prosthetic movement with respect to the bone which is to be detected.

## 2 Methods and Materials

All of the studies described here were approved by the Institutional Review Board and informed consent was obtained when required. All studies were performed using *routine* clinical procedures and scanners and were transferred via network in DICOM 3 (CT, MRI, PET) or in original format (SPECT) to a common computer system and converted to the standard AAPM format using Interformat [12].

All MRI breast scans were performed on one of five clinical scanners (Avanto, Symphony, TimTrio, Magnetom Vision Plus, and Sonata, Siemens Medical Solutions, Malvern, PA). All studies included an axial 3D T1-weighted gradient echo sequence performed before and three times after a rapid bolus injection of gadolinium contrast agent. These 3D volumes are reconstructed in the sagittal plane. An axially reconstructed 3D breath-hold fat-saturation sequence was available for 21 of the 23 patients; for the other 2 patients locally developed software was used to remap the volumes into the axial plane (with reduced resolution in the x direction). The MRI volumes were reconstructed into various matrix sizes which ranged in x from 160 to 512 pixels, in y from 320 to 512 pixels, and in z from 15 to 384 slices. The pixel sizes ranged in x from 0.625 - 1.6 mm, in y from 0.52 - 1.09 mm, and the slice thickness, z, ranged from 0.52 - 6.0 mm.

CTA cardiac studies were acquired on one of two MDCT scanners (Sensation 16 or Sensation 64, Siemens Medical Systems, Malvern, PA). Scanning was performed from the tracheal bifurcation to the diaphragm using a bolus tracking technique following the injection of 80-140 mL (80 for Sensation 64, and 140 for Sensation 16) of non-ionic contrast medium (Ultravist, Berlex, Montclaire, NJ). The gantry rotation time was 330 ms. Between 224 and 317 overlapping transaxial sections were reconstructed with a medium-sharp convolution kernel (B30f) into a 512x512x2 byte matrix with an x-y pixel size of 0.254-0.293 mm, a slice spacing of 0.5 mm, and an ECG-gated half-scan algorithm, with a resulting temporal resolution of 165 ms at the center of rotation. Volume reconstruction was retrospectively gated to the ECG. All CT orthopedic studies were acquired on the same scanner (GE Healthcare LightSpeed QX/i, Milwaukee, WI). Between 163 to 196 transaxial sections were reconstructed into 512x512x2 byte matrices with an x-y pixel size of 0.188-0.215 mm and a slice spacing of 0.7 mm.

All PET/CT scans were acquired on the same scanner (CTI HiRes/Biograph6, Siemens Medical Systems, Knoxville, TN). Forty-five minutes following the intravenous administration of 370 to 740 MBq (10 to 20 mCi) a routine whole body F-18 FDG scan was performed supine, followed by an additional prone scan from the base of the neck to the diaphragm. Emission scans were acquired following the acquisition of a low dose contrast-enhanced CT scan used for attenuation correction. Between 70-157 prone breast slices were reconstructed into 168x168x2 byte matrices, with CT-based attenuation correction. Only the attenuation-corrected data sets were used. The slice spacing was 2.5-4.0 mm and the x-y pixel size was 4.06 mm.

The SPECT scans were acquired on one of two scanners (ADAC Vertex or Cardio, Philips Medical Systems, Milpitas, CA). Using the clinical cardiac protocol, 61 projection slices gated over 17 time intervals (1037 slices) were reprocessed into 64 summed projection slices. These data sets were reconstructed into short- and long-axis volumes and, for this project, reconstructed into transverse slice volumes

consisting of 64x64x2 byte matrices per slice, with 16 to 23 slices each. The cubic voxels were 6.47 mm on each side. The commercially supplied reconstruction algorithm, which included a Butterworth filter (order:10, cutoff: 0.5 - rest; order: 5, cutoff: 0.66 - stress), was used.

### 2.1 Breast MRI/PET Studies

Data volumes were prospectively gathered from 23 women (ages 24-65 years) with suspected primary or recurrent breast cancer. 18-F FDG PET and MRI scanning was performed within the same day to four months of each other. The MRI breast scans were performed prone as they employ a special breast coil which allows the breasts to hang pendant; we designed and built an apparatus for the PET scanner which allows us to acquire the 18-F FDG PET scans with the patient in a similar position.

### 2.2 Cardiac CT/SPECT Studies

Data volumes were retrospectively gathered from seven patients (two women; five men; age range 47-72 years) with known or suspected heart disease. The studies were performed within 12 days to 11 months of each other.

### 2.3 Orthopedic Wrist Studies

Two human cadaver arms with a cemented implant (Universal Total Wrist, KMI, San Diego, CA) were used. In one arm, five 1 mm tantalum balls were implanted in the distal metaphysis of the radius where the bone/prosthesis ratio was smallest, and six in the carpal bones surrounding the distal component of the prosthesis. Five CT scans of each arm were obtained, with the arms repositioned at increasing angles to resemble five different patient positions. These ten CT scans were combined into case pairs, resulting in thirty different registration experiments defined by the placement of the landmarks (10 carpal bone only, 10 markers only, 10 carpal bone plus markers). This method was also applied to CT scans of a patient with symptomatic loosening of a total wrist implant. CT-guided volume registration showed that the etiology of the patient's pain was loosening of the prosthesis and a revision operation occurred.

### 2.4 Volume Fusion

A previously described [13-15] volume fusion tool co-developed with RAHD Oncology Products was used to register/merge each of these clinical studies. The registration algorithm incorporated in this tool requires the user to pick co-homologous points, i.e., corresponding point pairs (landmarks), on concurrently viewed slices that display the same physiologic point or structure. The transformation can be performed either using a first or second order polynomial (warping) or as a rigid body transformation requiring that as few as four point pairs be provided. Evaluation of the volume fusion was done visually in 2D and 3D [14], numerically (based on the distance between the actual position of the transformed landmarks and their ideal position (the reference landmark positions)), and through clinical

follow-up. Additional numerical tests, such as cross correlation of the volumes before and after registration and the squared sum of voxel value differences have also been done [14,17].

### 2.5 Automatic Registration

To test our results against automatic registrations, two programs based on mutual information (MI) or on normalized MI (NMI) were used. Program one offered an initial registration followed by a precise match based on MI [18], a surface match, a landmark match, and an option to adjust the match manually (3D Task card, Siemens Medical Solutions, Malvern, PA). Program two, developed by a research group and subsequently licensed commercially (Analyze 6.0, Mayo Clinic, Rochester, MI), provides three 3D methods: surface (which must be segmented for the brain), voxel intensity (NMI with extensive image/volume sub-sampling, and a novel gray scale binning) and non-rigid registration (iterative NMI). The National Library of Medicine (NLM) image tool kit (ITK) based in MI (which uses a gradient decent optimizer and [19]), is available in program two.

## 3 Results

In all cases, fusion using our program was accomplished within clinically acceptable ranges of registration quality, time from start to finish of the process, ease of use, and level of training necessary to operate the software. Multi-variant analysis performed on the transformed landmark distance differences showed no significant difference, confirming that the landmark choices between landmark sets and across studies (patients or cadaver arms) were consistent. A Student's t-test performed on these data showed in all cases that the transformed volume was at least within three reference pixels.

In the breast study, the MRI and PET scans were originally read separately. After fusion, a nuclear medicine physician and a radiologist re-read the studies, both side-by-side and fused. In the majority of cases, it was judged that the fusion increased the confidence of the readers and their clinical assessments and hence could in the future potentially spare patients unnecessary biopsies or anxiety-producing clinical follow-up. In this particular group of patients, pathology reports and more than six months of clinical follow-up were available. In this study, fusion showed that indeterminate enhanced lesions on MRI scans, when anatomically correlated with increased 18-F FDG uptake on PET, could be identified 98.6% of the time with cancer. Similarly, lack of such increased uptake could be identified 78.6% of the time as a benign process. Fig. 1 shows a cancer, (proven by biopsy); the original 18-F FDG PET scan (A), the MRI scan with enhancing lesion (B), and the matched PET superimposed upon the MRI scan (C). The fusion clearly demonstrates the location of the increased uptake. Also shown are the automatic matches obtained using both program one and two (voxel intensity only, neither of the other methods, surface, non-rigid gave a

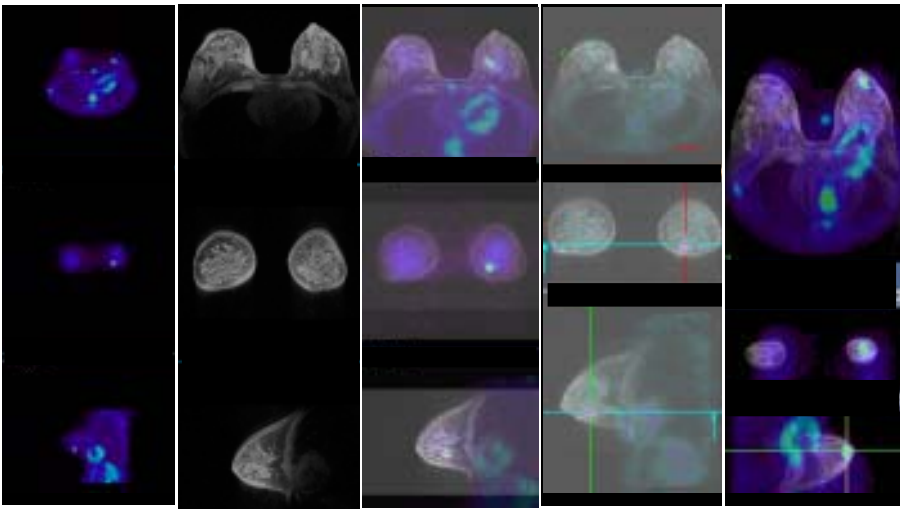



**Fig. 1.** Simultaneous display of axial (top) coronal (middle) and sagittal (bottom). (A) Uptake is visible on original 18-F FDG PET scan. (B) An enhancing lesion is seen on the original MRI scan. The PET and MRI are shown registered and superimposed using our method (C); using automatic method one (D); and using automatic method two (E).

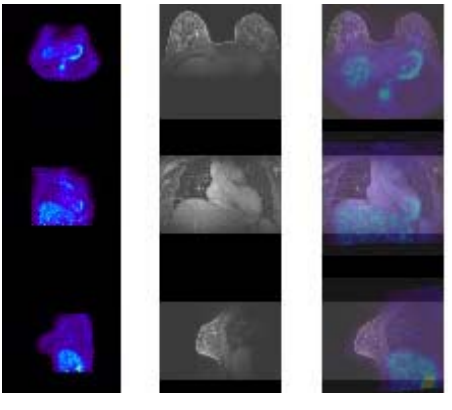



**Fig. 2.** Simultaneous display of axial (top) coronal (middle) and sagittal (bottom). There is no uptake is visible on original 18-F FDG PET scan (A).An enhancing lesion is seen on the original MRI scan (B). PET and MRI scans are shown registered and superimposed using our method (C).

**Fig. 3.** The segmented coronary artery tree has been superimposed on the registered SPECT heart which is shown translucently

result, and ITK gave a result similar to the above). Note also that program two displays the sagittal slice in the opposite direction from the other methods. Fig. 2 shows a biopsy-proven benign lesion (presented with the original unregistered scans and the two superimposed); note that this PET scan does not demonstrate significant uptake.

For the cardiac study, the CTA and SPECT scans were registered and viewed both in 2D and 3D. The coronary artery tree was segmented from the CTA, and the registered SPECT scan was displayed in 3D superimposed on it. If there was significant coronary artery narrowing, this would have been seen as diminished myocardial perfusion in the corresponding supplied territory on the SPECT volume. A 3D rendering of the registered SPECT heart with the coronary artery tree superimposed is shown in Fig. 3. A 2D view of the sliced volumes is shown in Fig. 4 presenting the original SPECT slices, the original CT slices, the SPECT volume transformed by our method and superimposed on the CT volume. The result of the match obtained using automatic program two is also shown (voxel intensity only - neither surface matching nor non-rigid registration worked; ITK gave very similar results to those shown). Unfortunately, we could not read the SPECT study into program one. Previously, we had a demonstration, using our cardiac data, of both program one and another (fusion 7D - Mirada) offered by the same company. The results were similar to those in Fig. 4D.

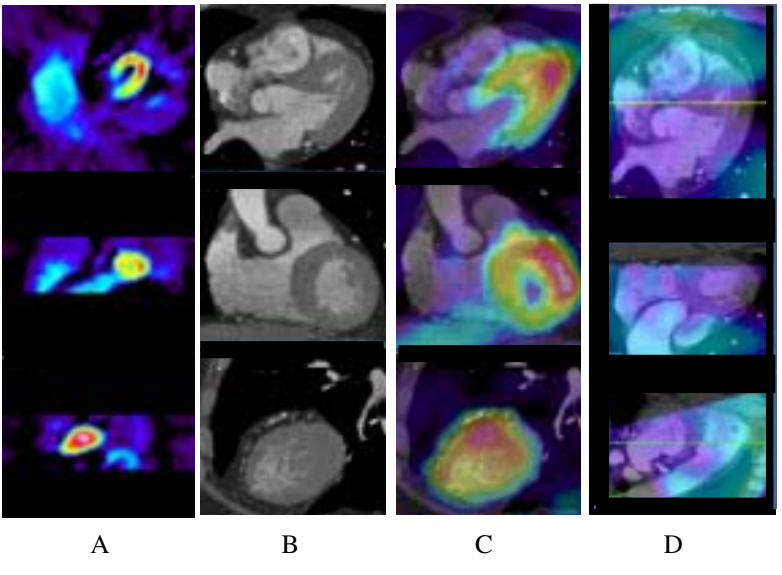



**Fig. 4.** Simultaneous display of axial (top) coronal (middle) and sagittal (bottom). (A) is the original SPECT scan. (B) is the original CTA scan. The SPECT and CT are shown registered and superimposed using our method (C); and using automatic method two (D).

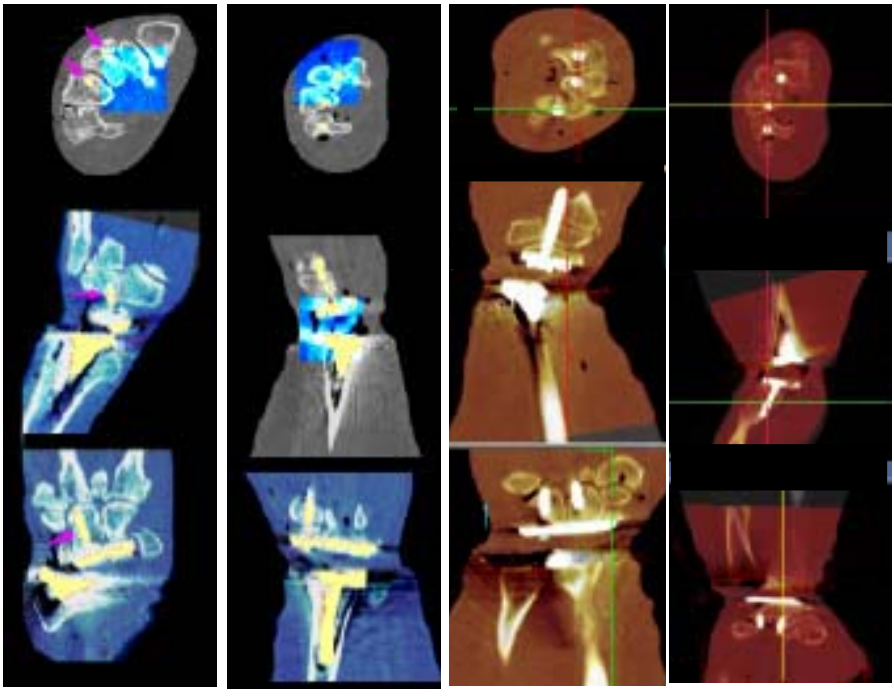



**Fig. 5.** Simultaneous display of axial (top) coronal (middle) and sagittal (bottom) from two different case matches (27 A and 1 B). The registration in A is based on both bone and tantalum markers. Arrows indicate the distal component of the prosthesis. The registration in B is based on bone alone. Both cases show well registered wrists using our method. The registration in C shows case 1 matched using automatic method one and the registration in D show the same case matched using automatic method two.

For the orthopedic study, visual and numerical methods confirmed that the fusion method could co-locate the wrist anatomy within 1 mm. Therefore, a displacement of greater than 1 mm could be detected, while standard radiographic methods can only detect a displacement of greater than 4 mm [20]. Fig. 5 demonstrates the visual quality of the match for two cases and shows the match from the automatic programs for case 1. Here the results for the automatic matches were quite comparable to ours, possibly because both images were of the same modality (CT). All automatic registration methods worked, except the for the non-rigid one.

## 4 Discussion

The results of breast MRI/PET fusion indicate that combining the functional tumor information provided by 18-F FDG PET scans together with the localization information from MRI scans can provide a better tool for diagnosis of primary and recurrent breast cancer. This could potentially save women with negative findings

from having to undergo the trauma of invasive or repeated procedures, as well as provide a more definitive diagnosis for affected patients. It could be argued that the CT scan associated with the PET scan could be used for anatomical localization of any 18-F FDG uptake. However, CT does not show the same subtle tissue differences as can be seen on MRI. In order to spare the patient unnecessary radiation, MRI is usually the first method of choice in screening patients with a family history of breast cancer, or in young patients with dense breasts. PET/CT is only used when the MRI scan is equivocal, which unfortunately has been reported to be as much as 50% of the time [8].

The results of the cardiac study identify the precise location and hemodynamic significance of obstructive coronary artery disease. By combining the anatomic coronary artery information obtained from CT volumetric data with the functional myocardial perfusion information obtained from SPECT data, a more comprehensive evaluation for coronary artery disease can be obtained. This information would be useful to clinicians by identifying patients with and without obstructive CAD. For patients without hemodynamically significant obstructive CAD, medical treatment of CAD can be initiated, and the patient could be spared the potential complications of invasive coronary angiography. In patients with evidence of significant obstructive CAD, the information obtained from the fused CT and SPECT volume sets can aid in procedural planning for either percutaneous or surgical revascularization.

The results of the wrist study demonstrate that CT image registration using this tool has the potential to detect prosthetic loosening at an earlier stage than is obtainable with planar radiography. This method is not yet as accurate as the radiostereometric analysis (RSA) technique [21]; however, (1) it is more readily available, (2) does not require the prior implantation of markers, (3) does not require as complex an analysis, (4) not only allows numerical analysis, but also gives a visual correlate to implant migration, and (5) is potentially lower in cost (although a detailed cost analysis has yet to be conducted).

These encouraging results are examples of what can be achieved when not constrained by the limitations of current automatic fusion tools. The landmark method initially may seem complex or time consuming, but the process is actually very straight forward and fast after a short learning curve. Once the user understands the process of picking landmarks for a modality and anatomical region, every new patient is done essentially the same way and the entire process generally takes ten to fifteen minutes. The learning curve is related to the image modality and anatomy, and involves learning to make some anatomical sense of each modality, developing a knowledge base of “dependable” anatomy in both modalities, and developing a procedure for using that information wisely. The main focus of the landmarks is in the region of critical interest, but by placing landmarks spread throughout the volume, a good registration can be obtained everywhere. The size of the data set is not as critical in determining time (as contrasted with some automatic tools [3,5]) as the modality and anatomical region.

## Acknowledgements

We wish to thank A. Deans, M. Kitazono, Y. Alcala, and J. Kardos for fusion, E.L. Kramer for contributions to the breast and heart projects, and C. Sollerman for contributions to the orthopedic project.

## References

1. Studholme, C., Hill, D.L.G., Hawkes, D.J. Automated 3D Registration of MR and CT Images of the Head. Medical Image Analysis **1** (1996) 163-167
2. Pluim, J.P., Maintz, J.B., Viergever, M.A. Mutual-information-based registration of medical images: a survey. IEEE Trans Med Imaging **TMI22** (2003) 986-1004
3. Meyer, C.R., Boes, J.L, Kim, B., Bland, P., Zasadny, K.R., Kison, P.V., Kora, K., Frey,K.A, and Wahl, R.L. Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin plate spline warped geometric deformations. Medical Image Analysis **3** (1996/7) 195-206
4. Knops, Z.F., Maintz, J.B., Viergever, M.A., Pluim, J.P. Normalized mutual information based registration using k-means clustering and shading correction. Medical Image Analysis 2005 (in press)
5. Shekhar, R., Walimbe, V., Raja, S., Zagrodsky, V., Kanvinde, M., Wu, G., Bybel, B. Automated 3-dimensional elastic registration of whole-body PET and CT from separate or combined scanners. J Nucl Med. **46** (2005)1488-96
6. Klabbers, B.A., deMunck, J.C., Slotman, B.J., Langendijk, H.A., deBree, R., Hoekstra, O.S.,Boellaard, R., and Lammertsma, A.A. Matching PET and CT scans of the head and neck area: Development and method of validation. Medical Physics **29** (2002) 2230-2238
7. Pfluger, T., Vollmar, C., WismŸller, A., Dresel, S., Berger, P., Suntheim, P., Leinsinger, G., Hahn, K. Quantitative Comparison of Automatic and Interactive Methods for MRI-SPECT Image Registration of the Brain Based on 3-Dimensional Calculation of Error. J. of Nuclear Medicine **41** (2000) 1823-1829
8. Liberman, L., Morris, E.A., Lee M.J., Kaplan, J.B., LaTrenta, L.R., Menell, J.H., Abramson, A.F., Dashnaw, S.M., Ballon, D.J., Dershaw, D.D. Breast lesions detected on MR imaging: features and positive predictive value. AJR **179** (2002) 171-178.
9. Hoffmann, U., Moselewski, F., Cury, R.C., Ferencik, M., Jang, I.K., Diaz, L.J., Abbara, S., Brady, T., Achenbach, S. Predictive value of 16-slice multidetector spiral computed tomography to detect significant obstructive coronary artery disease in patients at high risk for coronary artery disease: patient-versus segment-based analysis. Circulation. **110** (2004) 2638-2643
10. Adams, B.D. "Total Wrist Arthroplasty". Orthopedics **27** (2004) 278-284
11. Olivecrona, H., Weidenhielm, L., Olivecrona, L., Noz, M.E., Maguire Jr., G.Q., Zeleznik, M.P., Svensson, L., Jonson T. Spatial component position in total hip replacement. Accuracy and repeatability with a new CT method. Acta Radiologica **44** (2003) 84-91
12. Reddy, D.P., Maguire Jr., G.Q., Noz, M.E., and Kenny, R. Automating image format conversion - twelve years and twenty-five formats later. In: Lemke, H.U., Inamura, K., Jaffee, C.C., and Felix, F. (ed). Computer Assisted Radiology - CAR'93. Berlin, West Germany: Springer-Verlag; (1993) 253-258

13. Olivecrona, L., Crafoord, J., Olivecrona, H., Noz, M.E., Maguire Jr., G.Q., Zeleznik, M.P., Svensson, L. Weidenhielm, L. Acetabular component migration in total hip arthroplasty using CT and a semi-automated program for volume merging. Acta Radiologica, **43** (2002) 517-527
14. Noz, M.E., Maguire Jr., G.Q., Zeleznik, M.P., Kramer. E.L., Mahmoud F., Crafoord J. A versatile functional-anatomic image fusion method for volume data sets. Journal of Medical Systems **25** (2001) 297-307
15. Gorniak, R.J.T., Kramer, E.L., Maguire Jr., G.Q., Noz, M.E., Schettino, C.J., Zeleznik, M.P. Evaluation of a semiautomatic 3D fusion technique applied to molecular imaging and MRI brain/frame volume data sets. Journal of Medical Systems **27** (2003) 141-156
16. Maguire Jr., G.Q., Noz, M.E., Rusinek, H., Jaeger, J., Kramer, E.L., Sanger, J.J., Smith, G.: Graphics applied to image registration. IEEE Computer Graphics and Applications **CGA11** (1991) 20-29
17. Kimiaei, S., Noz, M.E. Jšnsson E., J. Crafoord J., and Maguire Jr., G.Q.: Evaluation of polynomial image deformation using anatomical landmarks for matching of 3D-abdominal MR images and for atlas construction. IEEE Trans Nuclear Science **NS46** (1999) 1110-1113
18. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., P. Suetens, P. Multimodality image registration by maximization of mutual information. IEEE Trans Med Imaging **TMI16** (1997) 187-198
19. Wells, W.M., Viola, P., Atsumi, H, Nakajima, S., Kikinis, R. Multi-modal volume registration by maximization of mutual information. Medical Image Analysis **1** (1996) 35-53
20. Malchau, H., KŠrrholm, J., Wang, Y.X., Herberts, P. Accuracy of migration analysis in hip arthroplasty. Digitized and conventional radiography, compared to radiostereometry in 51 patients. Acta Orthop Scand. **66** (1995), 418-424
21. Selvage, G. Roentgen stereo photog ram me try. A method for the study of the kinematics of the selectee system. (1974) Thesis, University of Lund, Lund, Sweden

# Comparison Between Parzen Window Interpolation and Generalised Partial Volume Estimation for Nonrigid Image Registration Using Mutual Information

Dirk Loeckx, Frederik Maes, Dirk Vandermeulen, and Paul Suetens

Medical Image Computing (ESAT/PSI), Faculty of Engineering
University Hospital Gasthuisberg, Herestraat 49, B-3000 Leuven, Belgium
Dirk.Loeckx@uz.kuleuven.ac.be

**Abstract.** Because of its robustness and accuracy for a variety of applications, either monomodal or multimodal, mutual information (MI) is a very popular similarity measure for (medical) image registration. Calculation of MI is based on the joint histogram of the two images to be registered, expressing the statistical relationship between image intensities at corresponding positions. However, the calculation of the joint histogram is not straightforward. The discrete nature of digital images, sampled as well in the intensity as in the spatial domain, impedes the exact calculation of the joint histogram. Moreover, during registration often an intensity will be sought at a non grid position of the floating image.

This article compares the robustness and accuracy of two common histogram estimators in the context of nonrigid multiresolution medical image registration: a Parzen window intensity interpolator (IIP) and generalised partial volume histogram estimation (GPV). Starting from the BrainWeb data and realistic deformation fields obtained from patient images, the experiments show that GPV is more robust, while IIP is more accurate. Using a combined approach, an average registration error of 0.12 mm for intramodal and 0.30 mm for intermodal registration is achieved.

## 1 Introduction

The goal of image registration is to find a transformation that maps positions of a reference image $I_R$ onto the corresponding positions of a floating image $I_F$ and is optimal in some sense. Different ways exist to judge the similarity between the reference and (deformed) floating image. They can be broadly classified into two categories: feature based and intensity based methods. In 1995, Collignon *et al.* [1] and Viola *et al.* [2] independently introduced mutual information (MI) as a similarity measure for intensity based medical image registration. Because of its robustness and accuracy for a variety of applications, either monomodal or multimodal, its popularity has been growing ever since [3, 4].

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 206–213, 2006.

Calculation of MI is based on the joint histogram of the two images to be registered, expressing the statistical relationship between image intensities at corresponding positions. However, the sought-after transformations will usually map pixels, located at integer positions in the reference image, to non integer locations in the floating image. This, together with the discrete nature of digital images, impedes the exact calculation of the joint histogram. In most cases, the intensities at the non integer locations are found by some kind of image interpolation in the floating image. Several reference articles have been published comparing different interpolators, yet no consensus exists on the best method [5, 6]. The joint histogram is usually constructed by grouping corresponding intensities in discrete, distinct bins [7, 5]. Alternatively, Thevenaz *et al.* [8] used a B-spline Parzen window approach to smooth the histogram bins.

As early as in 1997, Maes *et al.* [9] introduced partial volume distribution interpolation, an alternative method for the construction of the joint histogram. No interpolation is used to estimate the unknown intensities. Instead, each reference intensity is paired with the intensities of the voxels neighbouring the non integer location in the floating image. For each joint intensity pair, the histogram is updated with a partial hit, using a trilinear kernel to weight the contribution. Recently, Chen *et al.* [10] extended this approach to generalised partial volume estimation, using higher-order B-spline kernels for the weighting to reduce the artifacts.

Within this article, the B-spline Parzen window approach using intensity interpolation and the generalised partial volume estimation approach are compared using a single dataset, deformation model and optimisation algorithm. Both algorithms are implemented using an analytical expression for the derivatives. Nonrigid registration involves a huge number of degrees of freedom. To reach the optimum in an acceptable time-span, calculation of the derivative of the similarity criterion with respect to the deformation parameters is required. Although those derivatives can also be calculated numerically, as e.g. in [11], analytically calculated derivatives are less sensitive to noise and therefore might lead to a better registration.

This paper is organised as follows. In the methodology section, more details about both histogram estimators, the B-spline deformation mesh and the optimisation and validation algorithm are given. Next, the experimental setup and results are presented. We finish with a short discussion and some indications for future work.

## 2 Implementation

### 2.1 Histogram Estimation

For the histogram estimation, B-spline intensity interpolation (IIP) [8] and (B-spline) generalised partial volume estimation (GPV) [10] are compared. In both cases, a quadratic B-spline kernel is used, either for the image interpolation and Parzen window (IIP) or for the histogram distribution weights (GPV). The

**Table 1.** Multiresolution settings. The image scale gradually increases while the mesh refines. MCPs lists the number of mesh control points. The time and warping index $\varpi$ are averaged over the nine patients.

| *Stage* | *Scale* | $\Delta_{xyz}$ *(voxels)* | *MCPs* | *Time (mm:ss)* | $\varpi$ *(mm)* |
|---|---|---|---|---|---|
| | | | | 0:00 | 32.2 |
| 1 | 1/4 | 256 | 81 | 0:43 | 11.9 |
| 2 | 1/4 | 128 | 192 | 0:31 | 6.8 |
| 3 | 1/4 | 64 | 540 | 0:42 | 5.7 |
| 4 | 1/4 | 32 | 2400 | 0:52 | 4.6 |
| 5 | 1/2 | 32 | 2400 | 4:05 | 3.7 |
| 6 | 1/2 | 16 | 13608 | 4:43 | 2.3 |
| 7 | 1/1 | 16 | 13608 | 40:17 | 1.1 |
| 8 | 1/1 | 8 | 86700 | 31:22 | 0 |

mathematical complexity of both algorithms is comparable, requiring 9 B-spline evaluations for the histogram estimation and 27 more for the derivative.

### 2.2 B-Spline Deformation Mesh

The nonrigid deformation is modelled by a B-spline deformation mesh [11, 12]. A grid of mesh control points is positioned over the image. To model a more global deformation, the grid spacing is large, yielding a coarse mesh with few control points. A fine mesh has a small grid spacing and many control points, allowing a more local deformation. This approach allows a gradual refinement of the deformation mesh by decreasing the grid spacing.

### 2.3 Optimisation

A multiresolution optimisation approach is adopted, using 8 multiresolution stages. The multiresolution approach not only increases processing speed, by performing the initial calculations on a smaller image. Gradually decreasing the grid spacing will first recover more global deformations and progressively advance to finer deformations, thus avoiding local optima and creating a more realistic deformation field. In each stage, a limited memory quasi Newton optimiser [13] is used, especially designed for a large number of parameters. The exact derivative is calculated for each iteration. An overview of the registration settings per stage is given in Table 1.

### 2.4 Warping Index

The registration error is measured by the warping index $\varpi$ [8], which is the root mean square of the local registration error in each voxel, expressed in millimetres. The region of interest over which the warping index is calculated is limited to all voxels inside the brain. Registrations with a final warping index above 1 mm are considered as a failure.

## 3 Experiments

### 3.1 Setup

Accurate validation of nonrigid registration requires a set of reference and floating images pairs and, for each pair, the ground truth deformation. A registration is performed on each image pair. Comparison of the such obtained deformation fields with the ground truth yields an estimate of the accuracy of the registration algorithm. However, ground truth deformations for nonrigid registration are hard or impossible to obtain for real clinical cases. Therefore, we deduced realistic artificial deformation fields and corresponding images starting from the BrainWeb [14, 15] database and real patient images.

The BrainWeb database consists of simulated T1, T2 and PD brain images of $181 \times 217 \times 181$ mm with a 1 mm voxelsize in each dimension, calculated from a single phantom. Therefore, the images are a priori in perfect alignment. To obtain a set of 9 realistic deformation fields, the BrainWeb T1 image is registered to 9 patient T1 brain images with dimension $256 \times 182 \times 256$ and various voxelsizes of about $1 \times 1.2 \times 1\text{mm}^3$. The residual error or average distance of each registration step with respect to the final result, is listed in the final column of Table 1 and shown as a dotted line in Figure 1. The numbers shown are averaged over the 9 patients.

The obtained deformations are applied to the BrainWeb T1, T2 and PD images to produce deformed T1′, T2′ and PD′ images for each patient. This way, a set of deformed images with known deformation fields is obtained. Next, the original BrainWeb T1, T2 and PD images were registered to the deformed T1′, T2′ and PD′ BrainWeb images. Thus, for each patient data set, 9 nonrigid patient-BrainWeb validation registrations are performed: 3 intramodal and 6 intermodal. In total for the 9 patients, 9×9 registration results are obtained, of which 3×9 intramodal and 6×9 intermodal.

The registrations were performed on an AMD Opteron cluster; each registration running on a single processor.

### 3.2 Results

To distinguish between robustness (the ability to recover large deformations) and accuracy (the distance between the registration optimum and the true optimum), two sets of experiments were performed: starting from the initial position ($\varpi_0 = 32.2$ mm), and starting from the position obtained after 3 iteration stages ($\varpi_0 = 5.7$ mm). The experiments were performed using 32, 64, 128 and 256 bins. An overview of the error and failures for different numbers of bins is show in Table 2. 128 bins was found to be the best choice, slightly better than 64 bins and clearly superior to as well 32 as 256 bins.

Figure 1(a) displays the results using either IIP or GPV. The upper curves represent registrations starting at the initial position. The IIP approach leads in 74 out of 81 cases to a failure, whereas the GPV suffers from only 8 failures. Starting from the position obtained after 3 iteration stages (lower curves, 5.7 mm

**Table 2.** Number of failures and average warping error for different number of bins, for intensity interpolation (IIP) and generalis partial volume estimation (GPV), starting at multiresolution stage 0 ($IIP_0$, $GPV_0$) or 3 ($IIP_3$, $GPV_3$). Each entry shows the warping index $\varpi$ (in mm) and number of failures.

| | *32 bins* | *64 bins* | *128 bins* | *256 bins* |
|---|---|---|---|---|
| $IIP_0$ | 0.383 / 58 | 0.299 / 66 | 0.368 / 74 | / 81 |
| $GPV_0$ | 0.341 / 22 | 0.318 / 8 | 0.317 / 8 | 0.324 / 7 |
| $IIP_3$ | 0.280 / 0 | 0.226 / 0 | 0.212 / 0 | 0.306 / 1 |
| $GPV_0$ | 0.334 / 0 | 0.312 / 0 | 0.309 / 0 | 0.315 / 0 |

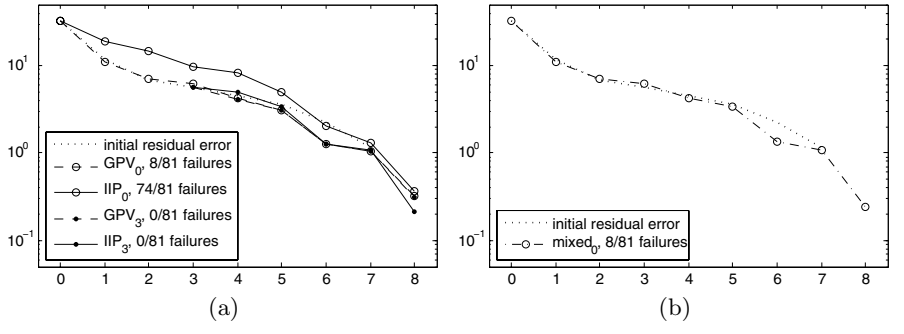



**Fig. 1.** (a) Warping index $\varpi$ (in mm) per multiresolution stage for IIP and GPV. $\varpi$ is always calculated compared to the final ground truth. The dotted line shows the residual error present in the different stages for the reference registration. The solid and dashed line show the error for intensity interpolation respectively generalised partial volume interpolation. The upper pair shows the results starting from the initial position, the lower pair starts from the third reference stage. (b) $\varpi$ per multiresolution stage for mixed interpolation (GPV for stage 1-4, IIP for stage 5-8).

average initial registration error), both approaches are always successful. The IIP approach now even outperforms the GPV approach with a final error of 0.21 mm compared to 0.31 mm. Closely comparing the lower GPV and IIP curve, it can be seen that while in stage 4 GPV is still superior, from stage 5 on the error decrease is larger for the IIP approach than for the GPV approach.

Therefore, we propose a mixed approach, using GPV for stage 1-4 and IIP for stage 5-8. The results are shown in Figure 1(b). As expected, the initial GPV stages brings the transformation sufficiently close to the optimum, with only 8 registration leading to a failure. The final IIP stages now are able to achieve a smaller registration error of 0.24 mm. Calculating the registration error independently for intermodal and intramodal registrations, gives an error (failure rate) of 0.30 mm (6/53) respectively 0.12 mm (2/27).

## 4 Discussion

Proper validation of nonrigid registration, especially at subvoxel accuracy, requires a pair of images for which the deformation is known exactly. In the approach we follow, this deformation is obtained by registering a base image to a patient image. Because we use the BrainWeb dataset, consisting of a priori registered T1, T2 and PD images, it enables the validation of multimodal registration.

In the validation approach we follow, the same deformation model is used to generate the ground truth and to perform the validation. We made this choice because we wanted to validate the quality of the similarity criterion and the optimiser rather than the applicability of the transformation model and possible regularisers. Validating the latter is much more cumbersome and application dependent and not the topic of this article. E.g., a different regulariser might be needed for inter-patient and intra-patient registration.

During registration, the algorithm evolves from an initial regime to a final regime. With ongoing multiresolution stages, the goal of the algorithm shifts from a coarse yet robust registration in the initial stages to an accurate registration in the final stages.

It is clear from Figure 1 that the IIP approach is insufficiently robust to cover the first registration stages. Figure 2 shows a typical trace of the mutual information around the origin and around the optimum of the first and fourth stage along the derivative. It shows that, in those stages, IIP is more sensitive to local optima and noisy. On the other hand, the warping index and Figure 1 indicate that, at the final stages, IIP is more accurate than GPV.

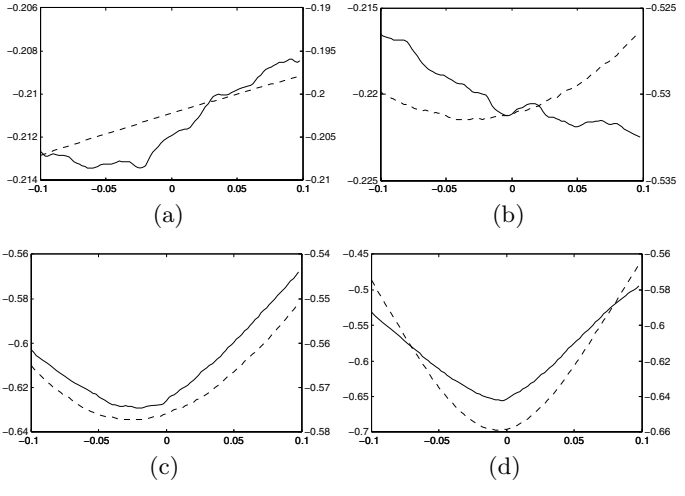
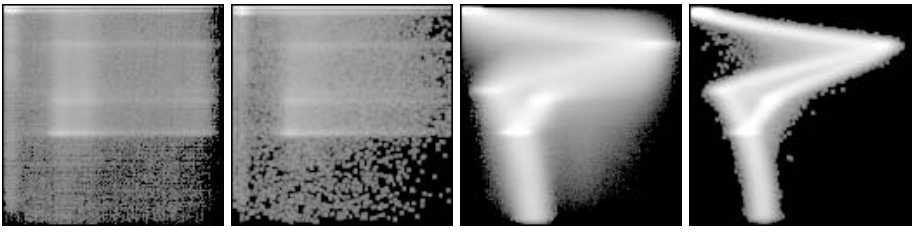



**Fig. 2.** Trace of the mutual information along the derivative (a, c) around the beginning and (b, d) around the optimum of the (a, b) first or (c, d) third stage. The solid line is for IIP, left Y-axis. The dashed line is for GPV, right Y-axis.



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

**Fig. 3.** (a,b) Initial (without deformation) and (c,d) final (at ground truth deformation) logarithm of the joint histogram for (a,c) GPV and (b,d) IIP interpolation

Thus, GPV interpolation is more robust, whereas IIP interpolation is more accurate. GPV will always spread the histogram update over the 3×3×3 neighbouring voxel intensities in the floating image. IIP will estimate a single intensity by image interpolation, and will spread the histogram update over the 3×3 neighbouring bins. The initial and final histogram for both IIP and GPV is show in Figure 3. Initially, it is more likely that the GPV approach will contribute, although partially, to a correct intensity pair. If this pair is updated repeatedly, its influence will become relevant. On the other hand, the IIP approach will create only a single joint intensity pair for every reference voxel, and thus is more influenced by local inhomogeneities or noise, leading to more spurious optima.

In the final registration stages, the distance between the estimated and true transformation is sub-voxel. Thus, the interpolated intensity of the IIP approach will be close to the true intensity. However, the GPV approach will still relate the reference intensity to a neighbourhood of 3×3×3 floating intensities. Therefore, the corresponding mutual information curve will be smoother and the location of the optimum less pronounced. The IIP approach, sufficiently close to the true optimum, will no longer be influenced by spurious optima.

## 5 Conclusion

The performance of intensity interpolation (IIP) and generalised partial volume distribution (GPV) for multiresolution nonrigid image registration were compared. Using the BrainWeb magnetic resonance data and realistic deformation fields, we have shown that for voxel based nonrigid image registration GPV is more robust, while IIP is more accurate. A combined approach using GPV for coarse and IIP for fine registration yields better registration results than an approach using only a single histogram estimator.

## Acknowledgements

This work is part of K.U.Leuven/OF/GOA/2004/05 and FWO G.0258.02.

## References


1. Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G.: Automated multi-modality image registration based on information theory. In Bizais, Y., Barillot, C., Di Paola, R., eds.: Proceedings XIVth international conference on information processing in medical imaging - IPMI95. Volume 3 of Computational Imaging and Vision., Kluwer Academic Publishers (1995) 263–274
2. Viola, P., Wells, W.M.: Alignment by maximization of mutual information. In: ICCV '95: Proceedings of the Fifth International Conference on Computer Vision, IEEE Computer Society (1995) 16–23
3. Studholme, C., Hill, D., Hawkes, D.: Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. Med Phys **24** (1997) 25–35
4. West, J., Fitzpatrick, J., Wang, M., Dawant, B., Jr., C.M., Kessler, R., Maciunas, R., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., van den Elsen, P., Napel, S., Sumanaweera, T., Harkness, B., Hemler, P., Hill, D., Hawkes, D., Studholme, C., Maintz, J., Viergever, M., Malandain, G., Pennec, X., Noz, M., Jr., G.M., Pollack, M., Pelizzari, C., Robb, R., Hanson, D., Woods, R.: Comparison and evaluation of retrospective intermodality brain image registration techniques. J Comput Assist Tomogr **21** (1997) 554–566
5. Lehmann, T., Gonner, C., Spitzer, K.: Survey: Interpolation methods in medical image processing. IEEE Trans. Med. Imag. **18** (1999) 1049–1075
6. Tsao, J.: Interpolation artifacts in multimodality image registration based on maximization of mutual information. IEEE Trans. Med. Imag. **22** (2003) 854–864
7. Pluim, J., Maintz, J., Viergever, M.: Mutual-information-based registration of medical images: A survey. IEEE Trans. Med. Imag. **22** (2003) 986–1004
8. Thevenaz, P., Unser, M.: Optimization of mutual information for multiresolution image registration. IEEE Trans. Signal Processing **9** (2000) 2083–2099
9. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imag. **16** (1997) 187–198
10. Chen, H.M., Varshney, P.K.: Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation. IEEE Trans. Med. Imag. **22** (2003) 1111–1119
11. Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D.: Nonrigid registration using free-form deformations: Application to breast MR images. IEEE Trans. Med. Imag. **18** (1999) 712–721
12. Loeckx, D., Maes, F., Vandermeulen, D., Suetens, P.: Non-rigid image registration using free-form deformations with a local rigidity constraint. In Barillot, C., Haynot, D.R., Hellier, P., eds.: Medical Image Computing and Computer-Assisted Intervention. Volume 3216 of Lecture Notes in Computer Science. (2004) 639–646
13. Byrd, R., Lu, P., Nocedal, J., Zhu, C.: A limited memory algorithm for bound constrained optimization. SIAM J. Sci. Comput. **16** (1995) 1190–1208
14. BrainWeb: `http://www.bic.mni.mcgill.ca/brainweb/` (1998)
15. Collins, D., Zijdenbos, A., Kollokian, V., Sled, J., Kabani, N., Holmes, C., A.C., E.: Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imag. **17** (1998) 463–468

# Elastic Registration Algorithm of Medical Images Based on Fuzzy Set

Xingang Liu and Wufan Chen

Institute of Medical Information, School of Biomedical Engineering
Southern Medical University, Guangzhou, China, 510515
{xjlxg, chenwf}@fimmu.com

**Abstract.** An elastic registration algorithm based on fuzzy set is proposed in the paper. The fuzziness of medical images is shown in two aspects: (1) the intensity of the pixels in medical images is fuzzy. The same kind of tissue may has different intensity and the same intensity may correspond to different tissues in one image; (2) the space position of image pixel is fuzzy. In the paper, we applied the fuzzy theory to the first aspect and presented the concept of fuzzy mutual information and its optimization method. For the second aspect, a multiresolution registration method based on fuzzy set and its optimization method is presented. 16 groups of experiments have been done and the results showed that the elastic registration algorithm based on fuzzy set can improve the accuracy and robustness of registration algorithm greatly.

## 1 Introduction

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. Its task is to find a correspondence function mapping coordinates from a reference image to coordinates of homologous points in a test image. Image registration is a crucial step in many image analysis tasks in which the final information is gained from the combination of various data sources like in image fusion, change detection, and multichannel image restoration. Image registration is mainly applied to the areas of video compression and coding, motion analysis and objection tracking, etc. It leads to algorithm for segmentation, general 3-D reconstruction, registration with annotated atlases, motion detection and compensation [1][2].

The character of fuzziness exists in medical images, which is shown in two aspects: one is the fuzziness of the pixel intensity in medical images. For the existence of imaging error and noise, one kind of tissue may has different intensity and the same intensity may correspond to different tissues in the same image, so it is necessary to apply the conception of fuzziness to imaging process. The other aspect is the space fuzziness of the pixels in medical images, which is mainly caused by various errors. The other reason for the space fuzziness is the application of multiresolution algorithm, in which a pixel in one image may correspond to different pixels in another image of different resolution. The fuzzy theory is applied widely in the field of image processing, including image segmentation[3]-[6], enhancement[7] , registration[8] and

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 214 – 221, 2006.

noise reduction[9]. The fuzzy theory and the conception of entropy are combined in [7], in which the maximum fuzzy entropy was applied to image segmentation. We combined the fuzzy theory [7][10] and mutual information (MI) together, and proposed the conception of fuzzy mutual information. We also constructed a multiresolution elastic registration algorithm based on fuzzy set, and compared it with a multiresolution elastic registration algorithm based on B-spline [11]. We applied fuzzy mutual information and multiresolution elastic registration algorithm to image registration and get a better result. The rest parts of the paper are arranged as the following: the tradition elastic registration algorithm based on mutual information is introduced in section 2; an elastic registration algorithm based on fuzzy set is introduced in section 3; the experimental results are shown in section 4; section 5 is the conclusion.

## 2 Tradition Elastic Registration Algorithm Based on MI

The input images are given as two m-dimensional discrete signals of the same size $f_y(i)$ and $f_b(i)$, where $i \in I \subset Z^2$, $I$ is the set of all pixel coordinates in the image. $f_y(i)$ and $f_b(i)$ are called the *reference* and *test* images, respectively. We suppose that the test image is a geometrically deformed version of the reference image. *g(x)* is a deformation (correspondence) function to be identified. The elastic deformation model is a hierarchical algorithm that adopts B-spline to construct the deformation function and interpolate the image[12].

### 2.1 The Estimation of Probability Density Function Based on Parzen Window

$L_t, L_r$ are the intensity sets of the reference and test images, respectively. The joint discrete Parzen histogram is defined as

$$h(l,k,C) = \frac{1}{h_e h_r} \sum_{x \in I} \sum_{y \in I} W\left(\frac{l - f_t(g(x,y,C))}{h_t}\right) \cdot W\left(\frac{k - f_r(x,y)}{h_r}\right) .$$

where $l \in L_t$, $k \in L_r$, $W(x)$ is the kernel function, and $\int W(x)dx = 1$.

Parzen probability is

$$p(l,k,C) = \alpha(C) \cdot h(l,k,C) .$$

where the normalization factor is

$$\alpha(C) = \frac{1}{\sum_{l \in L_t} \sum_{k \in L_r} h(l,k,C)} .$$

The MI of the reference and test images is

$$S = -\sum_{l \in L_t} \sum_{k \in L_r} p(l,k,C) \cdot \log \frac{p(l,k,C)}{p_t(l,C) \cdot p_r(k,C)} .$$

where $p_t(l,C) = \sum_{k \in L_r} p(l,k,C)$ and $p_r(k,C) = \sum_{l \in L_t} p(l,k,C)$

### 2.2 Optimization Method

In [12], B-spline is adopted to construct the deformation function and interpolate the images. The deformation function is

$$\begin{cases} g_x(x,y) = \sum_{i\in I_c}\sum_{j\in I_c} C_{x,i,j}\beta_n(x/2^w - i)\beta_n(y/2^w - j) \\ g_y(x,y) = \sum_{p\in I_c}\sum_{q\in I_c} C_{y,p,q}\beta_n(x/2^w - p)\beta_n(y/2^w - q) \end{cases}.$$

The interpolation function is

$$f^c(x,y) = \sum_{i\in I}\sum_{j\in I} b_{i,j}\beta_n(x-i)\beta_n(y-j).$$

We adopted the gradient descent method as the optimization method.

$$C_{i+1} = C_i - \mu\nabla_C S(C_i).$$

If we choose B-spline as the kernel function , the gradient of MI is

$$\nabla_C S(C) = \frac{\partial S}{\partial C} = -\sum_{l\in L_t}\sum_{k\in L_r}\frac{\partial p(l,k,C)}{\partial C}\log_2\left(\frac{p(l,k,C)}{p_t(l,C)}\right).$$

where

$$\frac{\partial p(l,k,C)}{\partial C_{x,i,j}} = \frac{1}{h_r h_t \|I\|}\sum_{x\in I}\sum_{y\in I}\beta_m\left(\frac{f_r(x,y)-k}{h_r}\right)\beta'\left(\frac{f_t(g_x(x,y),g_y(x,y))-l}{h_t}\right)$$
$$\cdot\frac{1}{h_t}\sum_{m\in I}\sum_{n\in I} b_{m,n}\beta'_m(g_x(x,y)-m)\beta_m(g_y(x,y)-n)$$
$$\cdot\beta_m(x/h-i)\beta_m(y/h-j)$$

## 3 Elastic Registration Algorithm Based on Fuzzy Set

### 3.1 Fuzzy Mutual Information Based on Intensity

#### 3.1.1 The Defination of Fuzzy Set

For $\sum_{i\in I}\beta_m(x-i)=1$ , we can choose B-spline as the membership function to implement the fuzzy division of the intensity space, so we select B-spline as the membership function of the fuzzy set.

#### 3.1.2 Fuzzy Probability Fuzzy Entropy and Fuzzy Mutual Information

We define the sets A and B as the fuzzy division sets of the intensity spaces of the reference and test images, which consist of fuzzy subsets $A_i$ and $B_i$, respectively. The definition for the sets A and B is the same. For each fuzzy set $A_i$ in $A$, the probability of the fuzzy sets $A_i$ is [10].

$$p(A_i) = \sum_{x \in A_i} \mu_{A_i}(x) p(x)$$

where $\mu_{A_i}$ is the membership function of fuzzy set $A_i$, and $\sum_{x \in A_i} \mu_{A_i}(x) = 1$. In the paper, we select B-spline $\beta_m$ as $\mu_{A_i}$.

The fuzzy entropy of the fuzzy set $A_i$ is

$$H(A_i) = p(A_i) \cdot \log p(A_i) = (\sum_{x \in A_i} \mu_{A_i}(x) p(x)) \cdot \log(\sum_{x \in A_i} \mu_{A_i}(x) p(x))$$

So the fuzzy entropy of the fuzzy division set A is

$$H(A) = \sum_{i=1}^{n} H(A_i)$$

The joint fuzzy entropy of the fuzzy sets A and B is

$$H(A,B) = \sum_{i=1}^{n} \sum_{j=1}^{m} p(A_i, B_j) \log p(A_i, B_j)$$

where we define

$$p(A_i, B_j) = \sum_{x \in A_i} \sum_{y \in B_j} \mu_{A_i, B_j}(x, y) p(x, y) = \sum_{x \in A_i} \sum_{y \in B_j} \mu_{A_i}(x, y) \mu_{B_j}(x, y) p(x, y)$$

### 3.1.3 Other Definition of Fuzzy Entropy

We define the fuzzy entropy of the fuzzy set $A_i$ as

$$H(A_i) = p(A_i) \cdot \log p(A_i) = (\sum_{x \in A_i} \mu_{A_i}(x) p(x)) \cdot \log(\sum_{x \in A_i} \mu_{A_i}(x) p(x)) \quad (1)$$

There are also many different definitions of the fuzzy entropy, the two main definitions are[7]:

$$H(A_i) = \sum_{x \in A_i} \mu_{A_i}(x) p(x) \log p(x) \quad (2)$$

$$H(A_i) = \sum_{x \in A_i} \mu_{A_i}(x) p(x) \log(\mu_{A_i}(x) p(x)) \quad (3)$$

In the three definition of the fuzzy entropy, our definition has two advantages: the first is the less computing time compared with the other two definitions, and there is only one operation of log in our definition; the second is the method defining the fuzzy entropy is the same with the tradition method defining the entropy, the only difference is the probability of the intensity is replaced by the probability of the fuzzy set $A_i$.

### 3.1.4 Optimization Method

Similar to the tradition optimization method of MI, we adopted B-spline to construct the deformation function and interpolate the images, and adopted the gradient descent method as the optimization method.

$$C_{i+1} = C_i - \mu \nabla_C S(C_i)$$

The gradient of the fuzzy MI is

$$\frac{\partial S}{\partial C} = -\sum_{A_i \in A} \sum_{B_i \in B} \frac{\partial p(A_i, B_i, C)}{\partial C} \log_2 \left( \frac{p(A_i, B_i, C)}{p_t(A_i, C)} \right)$$

where

$$\begin{aligned} \frac{\partial p(A_p, B_q, C)}{\partial C_{x,i,j}} = & \frac{1}{h_r h_t \|I\|} \sum_{k \in A_p} \sum_{l \in B_q} \mu_{A_p}(k,l) \cdot \mu_{B_q}(k,l) \\ & \cdot \sum_{x \in I} \sum_{y \in I} \beta_m \left( \frac{f_r(x,y) - k}{h_r} \right) \beta' \left( \frac{f_t(g_x(x,y), g_y(x,y)) - l}{h_t} \right) \\ & \cdot \frac{1}{h_t} \sum_{m \in I} \sum_{n \in I} b_{m,n} \beta'_m(g_x(x,y) - m) \beta_m(g_y(x,y) - n) \\ & \cdot \beta_m(x/h - i) \beta_m(y/h - j) \end{aligned}$$

### 3.2 Multiresolution Registration Algorithm Based on Fuzzy Set

Like section 3.1.1, we also select B-spline as the membership function to implement the fuzzy division of the coordinates spaces. After that, the fuzzy sets of pixels replace the pixels as the basic elements in the registration process. Various operations, such as the interpolation, constructing the similar criterion, or computing the gradient, are implemented on the basis of fuzzy sets. The operation details are similar to the ones based on the pixels, and the only difference is that the intensity of the pixels is replaced by the fuzzy intensity of the fuzzy sets. The fuzzy intensity of the fuzzy sets is

$$f(A_i) = \sum_{x \in A_i} \beta_m(x - i) f(x)$$

where x is the coordinates value of pixels, i is the index of the fuzzy sets. If we select the function $f(x) = \sum_{i \in I} \beta_m(x/2^w - i)$ as the membership function, with the increase of W ($W \in Z$), the range of the membership function increase and the number of the fuzzy sets decrease simultaneously, so we can complete the multiresolution registration algorithm based on the fuzzy sets. The optimization method is similar to the one of the tradition MI.

## 4 Experiment

This section presents a series of experiments in a controlled environment to assess the accuracy and robustness of our algorithm. We compare our test results with the ones of P. Th´evenaz 's elastic algorithm in [12],which is called reference algorithm in this paper. We select mutual information(MI) I(X;Y)=(H(X)+H(Y))/H(XY) as the similarity criterion. The size of original images is 256×256 pixels. We warp the images with the same deformation, which include the displacements of 10 pixels along X and Y coordinates respectively, rotation of 10 degrees anticlockwise and local deformation warped by the function 5×sin(x/128).

### 4.1 Results of Experiments

We selected 16 medical images arbitrarily from our image database as reference images, which involve 8 CT images and 8 MRI images. We stop when the step size $\Delta c$ falls below an *a priori* given threshold, which is set to 0.3 for all levels except the last, where we set it to 0.1. We do the experiments on the 16 groups of images by our and reference algorithms, respectively. The final results of the experiments evaluated by three experts in Nanfang hospital is shown in table 1.

**Table 1.** The comparison of 16 groups of test results

| item | CT image | MR image | Sum |
|---|---|---|---|
| A | 8 | 8 | 16 |
| B | 8 | 5 | 13 |
| C | 100% | 62.5% | 81.25% |
| D | 3 | 2 | 5 |
| E | 37.5% | 25% | 31.25% |

Notes: A:the sum of samples; B:the sample quantity of accurate matching by our algorithm;C:the accuracy rate by our algorithm; D:the sample quantity of accurate matching by the reference algorithm; E: the accuracy rate by the reference algorithm

In the experiments by our fuzzy algorithm, we achieve a total accuracy rate of 81.25% and even an accuracy rate of 100% for CT images, while the reference algorithm fails and only achieve a total accuracy rate of 31.25%. Our algorithm is superior to the reference one obviously under the same condition.

### 4.2 Analysis of Test Results

We select one group of MRI image to illustrate the results of registration. We yielded images of 128×128 pixels and 64×64 pixels respectively through the fuzzy multiresolution algorithm. The results of the experiment are shown in Fig.1.

All figures and table prove that the fuzzy algorithm can align not only the rotation and displacement deformation, but also the elastic one. The precision of registration is also up to sub-pixel level. The reference algorithm can align most of the deformation in the test image, but in some regions the misaligning between the test and reference images can be seen clearly.

### 4.3 Noise Dependence

We select 5 groups of images with accurate matching result as test samples, in which we added various levels of noise to demonstrate the influence of noise on the registration results. We set the variance of noise as 0.1 and 0.5, respectively. Under the noisy condition, our algorithm gets a better result than the reference one. The accuracy rate curves of the two algorithms are shown in Fig. 2.

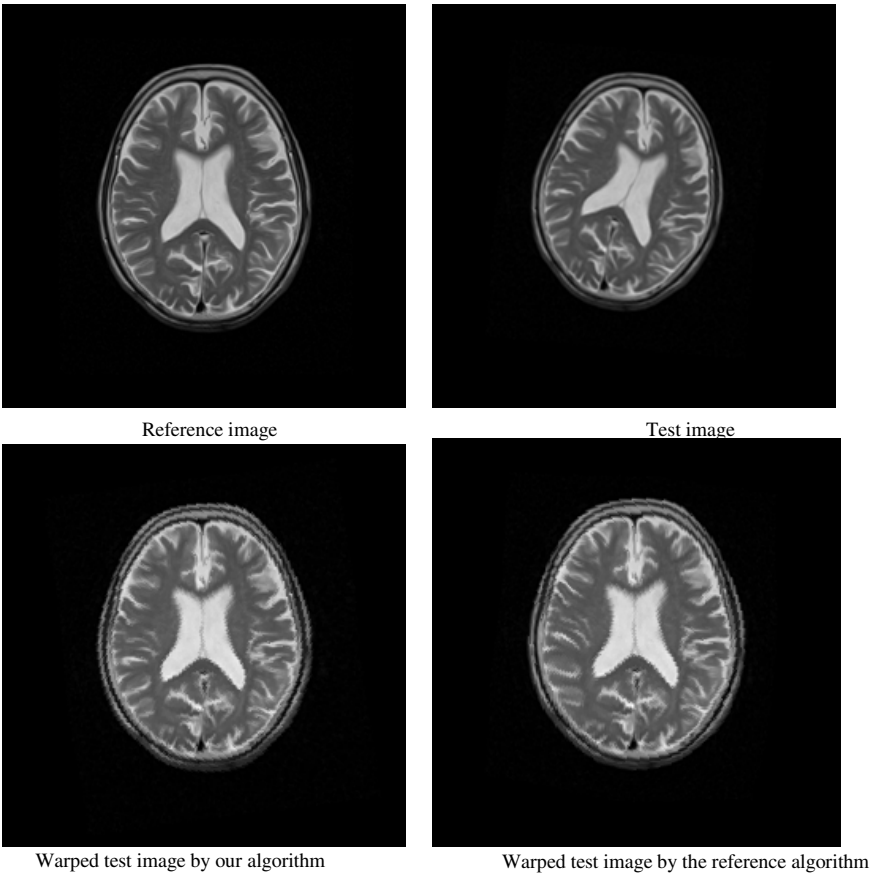



**Fig. 1.** The comparison of the experiment results by different algorithms

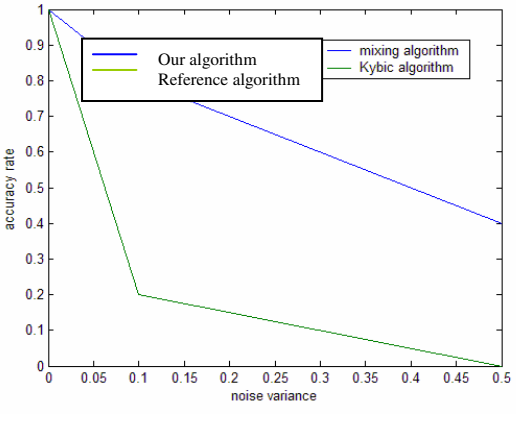



**Fig. 2.** The comparison of the accuracy rate with noise of different variance

## 5 Conclusion

We propose a fuzzy elastic registration algorithm that can align both rigid and elastic deformation. The fuzzy algorithm is more robust than the reference algorithm. Under the condition without noise, the fuzzy algorithm achieved a total accuracy rate of 81.25% and even 100% for CT images. Reference algorithm didn't get a satisfying result and only 5 groups succeed in matching. In the 5 groups of test with different level of noise, the accuracy rate by the reference algorithm decline more greatly than the fuzzy algorithm with the increase of noise variance. The experiments show that the fuzzy algorithm we proposed is more accurate and robust than the reference algorithm.

## Reference

[1] Barbara Zitova´, Jan Flusser . Image registration methods: a survey. Image and Vision Computing,2003,21:977-1000.
[2] J. B. Antoine Maintz, Max A. Viergever. A Survey of Medical Image registration. Medical Image Analysis, 1998, 2(1):1-37.
[3] Helene Caillol, Alain Hillion, Wojciech Pieczynski. Fuzzy Random Fields and Unsupervised Image Segmentation. IEEE Trans Geoscience and Remote Sensing, 1993, 31(4): 801-810.
[4] Helene Caillol, Wojciech Pieczynski, Alain Hillion, Estimation of Fuzzy Gaussian Mixture and Unsupervised Statistical Image Segmentation, IEEE Trans image procession, 1997, 6(3): 425-439.
[5] Dzung L. Pham, Jerry L. Prince. An Adaptive Fuzzy C-Means Algorithm for Image Segmentation in the Presence of Intensity Inhomogeneities. Pattern Recognition Letters, 1998, 20: 57-68.
[6] Dzung L. Pham, Jerry L. Prince. Adaptive Fuzzy Segmentation of Magnetic Resonance Image. IEEE Trans Medical Imaging, 1999, 18(9): 737-752.
[7] Cheng H D , Chen YH , Sun Y. A novel fuzzy entropy approach to image enhancement and thresholding. Signal Processing , 1999 , 75 (3) : 277 - 301.
[8] Amer Butt, Raj Acharya, Claudio Sibata, et al. Surface Matching of Multimodality Image Volumes By A Fuzzy Elastic Registraton Technique. Computerized Medical Imaging and Graphics, 1998,22:13-23.
[9] Dimitri Van De Ville, Mike Nachtegael, Etienne E. Kerre, et al. Noise Reduction by Fuzzy . Image Filtering. IEEE Trans Fuzzy System, 2003,11(4).
[10] Zadeh L A. Probability measures of fuzzy events. Math Anal Appl , 1968 , 23 : 421 - 427.
[11] Jan Kybic ,Michael Unser. Fast Parametric Elastic Image Registration. IEEE Trans Image Processing, 2003,12(11):1427-1442.
[12] P. Th´evenaz ,M. Unser. Optimization of mutual information for multiresolution image registration. IEEE Trans Image Processing, 2000, 9(12): 2083–2099.

# PET/CT Rigid-Body Registration in Radiation Treatment Planning Settings: Phantom Validation and Strategy Investigation

Sylvia J. Gong[1], Graeme J. O'Keefe[1,2], Kym Rykers[3], Carmel Mantle[3], Dishan H. Gunawardana[1], Aldo Rolfo[3], Morikatsu Wada[3], and Andrew M. Scott[1,2,4]

[1] Centre for PET,
[2] Department of Nuclear Medicine,
[3] Radiation Oncology Centre,
[4] Ludwig Institute for Cancer Research, Austin Health, Melbourne, VIC 3084, Australia
gong@petnm.unimelb.edu.au

**Abstract.** Therapeutic outcomes from radiation treatment planning (RTP) may be improved by combining tumour metabolism data with anatomic location via PET/CT image fusion. The aim of this work is to reveal the spatial accuracy, reproducibility and speed of registration techniques used in Gemini PET/CT system as a part of quality assurance. A multi-layer alignment device and a Rando Man phantom were used to validate PET/CT intrinsic alignment. Known transformations were performed to simulate different misalignment between PET and CT phantom images. Point-based and image-based registration techniques used to correct misalignment were assessed and compared quantitatively by measuring absolute distances between the centroids of corresponding fiducial markers in the registered volumes. In comparison with phantom studies using image-based registration, point-based registration shows better spatial accuracy and faster correlation, but longer time and more effort in human intervention. The experimental findings confirm that Gemini can produce robust intrinsic image alignment, as well as the accurate PET/CT image registration with careful user interaction.

## 1 Introduction

Integrated with powerful digital techniques, positron emission tomography (PET) is evolving as the most rapidly growing functional imaging modality in nuclear medicine and oncology. The high accuracy of PET in cancer detection and staging has major implications in terms of improving the treatment planning and avoiding unnecessary treatment and its associated morbidity and cost[1].

PET scans alone, however, are usually difficult to interpret properly in the absence of precise anatomical correlation [3,4]. In spite of the improved specificity and sensitivity of tumor assessment by using anatomical images side-by-side[4] or registering PET images with CT[5,6], PET images display few anatomical landmarks thus preventing direct correlation with structural images[2]. Registration uncertainty also occurs due to device, protocol and time-point differences in data acquisition with patient repositioning, patient scheduling and possible involuntary and incontrollable organ motion and inter-section anatomical changes[3,6]. With the recent advent of hybrid PET/CT

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 222 – 229, 2006.

units[3,7] which can perform both CT and PET in a single imaging session and produce fast intrinsic-fusion of CT and PET images with minimal user interaction, these issues should be one step closer to a solution. Because of the exact anatomic correlation of metabolic abnormalities, the resulting improvement in precision and accuracy of in-line PET/CT[8] offers a number of advantages to augment the treatment planning of head-and-neck cancer, including improved confidence in image interpretation, reduced risk of tumor-tissue misses, minimized dose of radiation to normal tissues, improved volume delineation as a result of decreased inter-observer variability, and avoidance of ineffective radiation therapy[9, 10].

The accurate multi-modality image registration is essential for the determination of the utility and impact of PET in the radiotherapeutic management of head and neck cancers. The registration procedure includes the registration of functional PET images with anatomical CT images at the site of PET Centre; and the registration of anatomical CT images with the co-ordinate system of planning CT scans used for RTP at the site of Radiation Oncology Centre. The registration process, however, is prone to errors and mis-registration could impact on the accuracy of target tumor volume delineation in RTP. This article presents ad hoc validations and results in verifying the reproducibility of intrinsic alignment, as well as in comparing and evaluating the spatial accuracy and speed of extrinsic PET/CT registration techniques, as a part of quality assurance at the site of PET/CT.

## 2 Materials and Methods

### 2.1 Phantoms with Fiducial Markers

A multi-layer fixture bearing six non-collinear 10mCi $^{22}$Na point sources was supplied by Philips for the use of image alignment calibration and verification. The fixture was also employed to perform validation experiments of PET/CT registration using the point sources as fiducial markers.

A partial Alderson Rando Man Phantom (head through mid-femur) was used to simulate more complex structure of human bodies. Nine metal cannulas, each of which is 1.2 mm in internal diameter and about 8 mm long, were used as external fiducial markers affixed to the landmarks throughout the surface of the Rando phantom. The cannulas containing $^{18}$F-FDG with concentration of 1MBq/mL can be easily recognized in both CT and PET emission images. To prevent the leakage of radioactivity while allowing the reuse of the cannulas, one end of the cannula was properly clamped and the other end remaining open. The opening end can be sealed properly with the plastic tape once filled with the radioactivity.

### 2.2 Acquisition Protocols

Diagnostic PET and low-dose CT scans were performed using a Gemini$^{TM}$ open system (Philips Medical Systems Inc., Cleveland, OH, USA) which integrates a Philips Mx8000 EXP Dual-slice Helical CT Imaging System and a Philips-ADAC Allegro$^{TM}$ full-ring 3D PET Imaging System with Pixelar$^{TM}$ GSO crystal detectors. An initial CT "Surview" scan was performed to determine the exact axial examination range followed by a non-contrast CT scan using the typical scan parameters: 3.2 mm slice

thickness with 3.2 mm spacing, 600 mm displayed field of view (FOV) with a 512x512 matrix, peak voltage of 120 kVp, 30 mAs per slice, spiral pitch of 1.5 and tube rotating time of 0.75 s. Once the acquisition was completed, the reconstructed CT image was automatically transferred from the acquisition computer to the PET reconstruction workstation in both DICOM and PETview™ formats for registration.

PET emission scans were acquired using 576 FOV, 4 minutes per bed position (1 bed position for the alignment fixture and 11 bed positions for the Rando phantom), and reconstructed using RAMLA[11]. Transmission scans acquired using 740MBq $^{137}$Cs transmission source were used for non-uniform attenuation and scatter correction of emission sinograms. Transmission images were also applied to correlate to CT using image-based registration method[12].

### 2.3 Phantom Experiments

We performed four validation studies to test the accuracy and coherence of intrinsic PET/CT coordinates prior and posterior to a series of clinical scans during a 6-month period, in which the multi-layer fixture and Rando phantom were scanned individually with both modalities in a single imaging session. The reconstructed CT and PET images were fused directly without any interference.

PET/CT image sets of each phantom acquired in the validation studies were artificially misaligned by applying ten pre-defined transformations on PET transmission and emission images (see Figure 1), five with translation only and five combined with translation and rotation. The resulting images were re-sampled and saved as the simulated images.

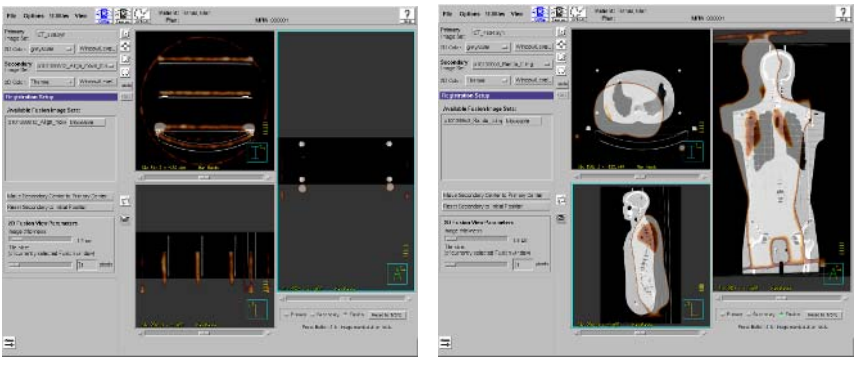


**Fig. 1.** Simulated Misalignment of PET Transmission and CT Images (left: translation of the alignment phantom image; right: translation and rotation of the Rando phantom image)

### 2.4 Image Registration

All CT and PET image pairs were imported in Syntegra™, a Philips built-in software package with Gemini for registration and fusion. Syntegra allows manual or semi-automated PET/CT image registration using a rigid-body transformation of linear and

angular position. It also enables visualizing registered image volumes in either independent or fusion mode.

Theoretically a sequence of rigid-body transformations can be used to map a given point in an active image volume (A) to a stationary reference image volume (B): system misalignment between the two coordinate systems, local transformation in the coordinate system of the active volume, and translation between the active image origin and the coordinate system origin of the active volume. The sequence of transformation to achieve the active/reference volume registration can be expresses as the combination of local and global transformation:

$$\widetilde{P}_{B} = R_{s\gamma} * R_{s\beta} * R_{s\alpha} * (\widetilde{P}_{Local} - \widetilde{T}_{O}) + \widetilde{T}_{O} + \widetilde{T}_{S} \tag{1}$$

$$\widetilde{P}_{Local} = R_{\gamma} * R_{\beta} * R_{\alpha} * (\widetilde{P}_{A} - \widetilde{T}_{O}) + \widetilde{T}_{O} + \widetilde{T} \tag{2}$$

where, $\widetilde{P}_{B}$ is the position vector of a given point in the coordinate system of the reference volume and $\widetilde{P}_{A}$ is the position vectors of the same point in the coordinate system of the active volume, respectively; $\widetilde{P}_{Local}$ is the position vector in the domain of the active volume after the local transformation about the active image origin. $\widetilde{T}_{S}$, $\widetilde{T}_{O}$ and $\widetilde{T}$ are the translation vectors of active/reference system offsets, active origin/image iso-centre offsets and absolute active image adjustment in x-, y- and z-axis, individually. $R_{\alpha}$, $R_{\beta}$ and $R_{\gamma}$ are the local tilt, yaw and roll matrices of the active volume; and $R_{s\alpha}$, $R_{s\beta}$ and $R_{s\gamma}$ are the active/reference system tilt, yaw and roll matrices, respectively.

### 2.4.1 Point-Based Registration

In Syntegra, an active and its reference image can be correlated by adjustingg transformation parameters of registration system directly. A movable active image can also be shifted and rotated manually to align with its reference image using 2D/3D graphic tools. In addition to these two interactive methods, two semi-automated image registration approaches are also provided: point-based and image based registration.

Point-based registration starts with the manual localization of fiducial points (FP) in CT and PET emission images, followed by an automated algorithm which generates a rigid-body transformation producing the least-square fit of the selected FP pairs[13]. At least three non-collinear FP pairs are needed to establish a unique transformation between two images.

Fiducial localization involves manual determination of intensity-weighted centroids of fiducial markers in PET images and geometrical centroids of fiducial markers in CT images using a cross-hair pointer in Syntegra with careful visual inspection. The selected point is correlated in trans-axial, coronal and sagittal 2D views which can compensate for parallax error by altering viewpoints. The displacements about the origin of 3D Cartesian image coordinates were also displayed in each window.

Although manual fiducial indication is labour-intensive, the more fiducial points that are used, the higher accuracy the registration can achieve[14]. To compromise the time used to select and define FPs and reliability of the algorithm to find an optimized transformation, 3 fiducial pairs were selected for the multi-layer phantom images and 6 fiducial pairs for Rando phantom images. The other FPs not used for the registration algorithm were also inspected after the registration to assess target registration error.

### 2.4.2 Image-Based Registration

Image-based registration approach in Syntegra employs similarity measures of relevant image segments to seek an optimized global transformation. We utilized the algorithm maximizing the normalized mutual information[15] for PET/CT registration as recommended[16]. In the registration, PET transmission images with much more anatomical information than emission images were correlated to CT images first. The resulting transformations were applied to PET emission images for quantitative assessment of registration errors at fiducial markers.

### 2.5 Visual Assessment and Quantitative Analysis

Registration errors were measured quantitatively as the mean distances between the corresponding FP sets, since the co-location between the FPs in the reference and correlated image volumes indicates a correct registration. Appropriate fiducial localization is essential for the accurate assessment of all registration techniques applied. The visual technique in which the correlated volumes displayed and inspected in fusion mode was used in all procedures involving manual fiducial localization.

## 3 Results

Table 1 shows the validation results of PET/CT intrinsic alignment. The table lists the mean and standard deviation of misalignment at 6 FP pairs in multi-layer phantom images and 9 FP pairs in Rando phantom. The misalignment (< 1mm) of four alignment validation studies shown in Table 1 verify that intrinsic PET/CT image pairs can be correlated accurately and robustly with no need of any further registration.

The results of point-based and image-based registration are showed in Table 2. The mean and standard deviation of registration errors in 5 different simulations for each group of data are listed. In comparison with studies using $^{22}$Na point sources, image registration of studies using metal fiducial markers shows better accuracy, by either point-based or image-based registration. Generally point-based registration shows better spatial accuracy (< 1mm) than the image-based registration (< 4mm).

**Table 1.** Validation of PET/CT intrinsic alignment

| Image Sets with Intrinsic Alignment | | | Mis-alignment (mm) | | |
|---|---|---|---|---|---|
| | | | x ± σ | y ± σ | z ± σ |
| Cross Fiducial Localization | multi-layer (6 fiducial points) | pre | 0.00 ± 0.48 | 0.79 ± 0.47 | 0.00 ± 0.00 |
| | | post | 0.01 ± 0.01 | 0.59 ± 0.64 | 0.00 ± 0.00 |
| | Rando (9 fiducial points) | pre | 0.00 ± 0.01 | 0.00 ± 0.00 | 0.00 ± 0.00 |
| | | post | 0.01 ± 0.01 | 0.00 ± 0.01 | 0.00 ± 0.00 |

**Table 2.** Registration errors in point-based and imaged-based registration

| Image Sets with Simulated Misalignment | | | Registration Error (mm) | | |
|---|---|---|---|---|---|
| | | | x ± σ | y ± σ | z ± σ |
| point-based image registration | multi-layer (3 fiducial pairs) | trans. | 0.22 ± 0.09 | 0.19 ± 0.12 | 0.07 ± 0.02 |
| | | trans. & rot. | 0.14 ± 0.13 | 0.33 ± 0.11 | 0.09 ± 0.22 |
| | Rando (6 fiducial pairs) | trans. | 0.10 ± 0.06 | 0.09 ± 0.05 | 0.05 ± 0.06 |
| | | trans. & rot. | 0.08 ± 0.06 | 0.11 ± 0.07 | 0.05 ± 0.05 |
| image-based image registration | multi-layer (without sub-volume) | trans. | 2.25 ± 0.59 | 2.48 ± 0.11 | 2.14 ± 0.34 |
| | | trans. & rot. | 3.18 ± 0.97 | 2.41 ± 0.37 | 2.64 ± 0.31 |
| | Rando (without sub-volume) | trans. | 1.28 ± 0.67 | 1.55 ± 0.30 | 1.78 ± 2.01 |
| | | trans. & rot. | 1.56 ± 0.60 | 1.77 ± 0.47 | 1.87 ± 1.85 |

Table 3 shows the processing time used for point-based and image-based registration procedures respectively. Although the processing time of point-based registration (< 1s) was much faster than the image-based registration (> 30s), the total time to complete a point-based registration was longer than the image-based registration due to the manual determination of fiducial points.

**Table 3.** Processing time used for the point-based and image-based registration procedures

| Image Sets with Simulated Misalignment | | | Processing Time | | |
|---|---|---|---|---|---|
| | | | manual (minute) | automated (second) | total (minute) |
| Point-base Registration | multi-layer (3 fiducial point pairs) | trans. | > 5 | < 0.5 | >5 |
| | | trans. & rot. | | <0.5 | >5 |
| | Rando (6 fiducial point pairs) | trans. | > 15 | < 0.5 | >15 |
| | | trans. & rot. | | <0.5 | >15 |
| Image-based Registration | multi-layer (without sub-volume) | trans. | 0 | 30.17±8.28 | < 1 |
| | | trans. & rot. | 0 | 33.08±11.18 | <1 |
| | Rando (without sub-volume) | trans. | 0 | 164.06±10.89 | <5 |
| | | trans. & rot. | 0 | 179.98±63.91 | <5 |

## 4 Discussion

Many PET/CT clinical and research applications require a very high degree of spatial accuracy of image registration. In this work we demonstrate a series of phantom experiments and simulation, as well as comparison and evaluation the spatial accuracy of different techniques used for PET/CT rigid-body image registration.

The accuracy of point-based image registration depends on the proper indication of non-collinear corresponding fiducial points. Our experience confirms that fiducial localization is labor-intensive and prone to error[17]. To minimize the observer variability, all fiducial localizations were repeated at least once. Any difference was reviewed and corrected using a final agreement. We find that cross visualization of PET and CT image volumes in fusion mode is helpful in localizing and identifying misalignment.

In comparison with studies using image-based registration, point-based registration shows better spatial accuracy (see Table 2) and faster correlation, but longer time and more effort in human intervention (see Table 3). Combining image- and point-based procedures to register image sets will be helpful in achieving optimal accuracy and speed.

Generally in registration of physically misaligned images, only the subject is considered as vital and can be manually segmented from patient bed in both CT and PET images using some sub-volume tool. The sub-volume selection was not performed in the registration of the simulated image sets reported here. The good results of the simulated data registered using normalized mutual information algorithm may be caused by the higher lever of similarity between scan volumes, in which the upper-pallet was reoriented together with phantoms using a global transformation.

Our validation study used the image sets of phantoms with fiducial markers and its simulation, of which the true transformation is known *a priori*. Any registration approach applied on the misaligned images can be evaluated using this "gold standard". In clinical practice, however, the correlation truth of human subject to test the registration against is rarely known. Assessment results may be limited by the reference method rather than the registration method being tested[18]. Further assessment of image registration on clinical studies can be verified using consistency test[2] alternatively.

## 5 Conclusion

The theoretical and experimental findings so far prove that Gemini can perform robust intrinsic system alignment with a spatial accuracy less than 1 mm. The semi-automated techniques in Syntegra can produce accurate PET/CT image registration with careful user interaction. The spatial accuracy of point-based registration is less than 2 mm and is less than 4mm in image-based registration. Overall the spatial accuracy of image registration is better than the spatial resolution of the PET scanner used (4.5 mm - 5 mm). This performance is sufficiently accurate for most PET/CT applications. Our experiences of comparing and evaluating the image display and processing also confirm that fiducial indication in fused PET/CT is superior to visual correlation of PET and CT.

## Acknowledgments

We thank Daniel Gagnon, Philips Medical Systems, for his valuable support.

## References

[1] AM Scott, “Current status of positron emission tomography in oncology,” *Australasian Radiology* 46: 154-162, 2002.
[2] BM Klabbers, JC de Munck, and BJ Slotman, et al., "Matching PET and CT scans of the head and neck area: Development of method and validation," *Medical Physics* 29: 2230-8, 2002.
[3] T Beyer, DW Townsend, and TM Blodgett, "Dual-modality PET/CT tomography for clinical oncology," *Quarterly Journal of Nuclear Medicine* 46: 24-34, 2002.
[4] P Reinartz, F Wieres, and W Schneider, et al., "Side-by-side reading of PET and CT scans in oncology: which patients might profit from integrated PET/CT?" *Eur J Nucl Med Mol Imaging* 31: 1456-61, 2004.
[5] RL Wahl, "Why nearly all PET of abdominal and pelvic cancers will be performed as PET/CT?" *JNM* 45:82S-95S, 2004.
[6] GW Goerres, GK von Schulthess, and HC Steinert, "Why most PET of lung and head-and-neck cancer will be PET /CT?" *JNM* 45: 66S-71S, 2004.
[7] DE Heron, RS Andrade, and J Flickinger, et al., "Hybrid PET-CT simulation for RTP in head-and-neck cancers: A brief technical report," *International Journal of Radiation Oncology, Biology, and Physics* 60: 1419-1424, 2004.
[8] C Cohade and RL Wahl, "Applications of positron emission tomography/computed tomography image fusion in clinical positron emission tomography-clinical use, interpretation methods, diagnostic improvements," *Seminars in Nuclear Medicine* 33: 228-37, 2003.
[9] Goerres, G. W., G. K. von Schulthess, et al. (2004). "Why most PET of lung and head-and-neck cancer will be PET/CT." *Journal of Nuclear Medicine* **45**(Suppl 1): 66S-71S.
[10] Ciernik, I. F., E. Dizendorf, et al. (2003). "Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study." International *Journal of Radiation Oncology, Biology, Physics* **57**(3): 853-63.
[11] J Browne and AB de Pierro, "A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography," *IEEE Trans. Med. Imag.* 15: 687-99, 1996.
[12] D Dey, PJ Slomka, LJ Hahn, and R Kloiber, "Automatic 3D multimodality registration using radionuclide transmission CT attenuation maps: a phantom study," *JNM* 40: 448-55, 1999.
[13] JM Fitzpatrick and JB West, "The distribution of target registration error in rigid-body point-based registration," *IEEE Trans. Med. Imag.* 20: 917-27, 2001.
[14] G Ende, H Treuer, and R Boesecke, "Optimization and evaluation of landmark-based image correlation," *Physics in Medicine & Biology* 37: 261-71, 1992.
[15] P Viola and WM Wells, "Alignment by maximization of mutual information," Proceedings of 5th International Conference on Computer Vision: 16-23, 1995.
[16] Philips Medical Systems (Cleveland), "Gemini Imaging System User's Manual Vol. 3 - PET Image Processing," 2003.
[17] CR Maurer (Jr.), JM Fitzpatrick, and MY Wang, et al., "Registration of head volume images using implantable fiducial markers," *IEEE Trans. Med. Imag.* 16: 447-62, 1997.
[18] JL Andersson, A Sundin, and S Valind, "A method for coregistration of PET and MR brain images," *JNM* 36: 1307-15, 1995.

# 3D Mouse Brain Reconstruction from Histology Using a Coarse-to-Fine Approach

Paul A. Yushkevich[1], Brian B. Avants[1], Lydia Ng[2], Michael Hawrylycz[2], Pablo D. Burstein[1], Hui Zhang[1], and James C. Gee[1]

[1] Penn Image Computing and Science Laboratory, Department of Radiology, University of Pennsylvania, Philadelphia PA, USA
[2] Allen Institute for Brain Science, Seattle WA, USA

**Abstract.** The Allen Brain Atlas project aims to bridge the divide between genomics and neuroanatomy by mapping the expression of the entire C57BL/6J mouse genome onto a high-resolution 3D anatomical reference atlas of the mouse brain. We present the image registration approach used to generate this anatomical reference from histological sections. Due to the large number of sections (525) and the presence of debris and distortions, a straightforward alignment of each slice to its neighbors fails to accurately recover the 3D shape of the brain. On the other hand, multimodality registration of histology slices to an MRI reference compromises correspondences between neighboring slices. Our approach combines the high-frequency component of slice-to-slice histology registration with the low-frequency component of the histology-to-MRI registration to produce a coarse-to-fine reconstruction that is accurate both in its global shape and in the alignment of local features.

## 1 Introduction

The problem of reconstructing a 3D volume from histological sections arises frequently in animal model research. In real experimental data, the reconstruction problem is made difficult by high incidence of artifacts such as tearing of tissue and debris. We present a new coarse-to-fine reconstruction algorithm that combines a graph-theoretic slice-to-slice reconstruction with a global histology-to-MRI reconstruction to achieve high accuracy both in the alignment of features between slices and in the 3D shape of the reconstructed brain. Our algorithm was used to generate the anatomical reference atlas of the mouse brain for the Allen Brain Atlas (ABA) project at the Allen Institute for Brain Science (AIBS).

The goal of the ABA project is to generate a freely accessible database of the expression of some 24,000 genes that compose the mouse genome. High resolution *in situ* hybridization in coronal and sagittal slices is used to map gene expression onto the reference atlas, allowing researchers to make complex queries relating gene expression and neuroanatomy [1]. Already, over 50% of the genome has been mapped and is available to researchers at `www.brain-map.org`.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 230–237, 2006.

## 2 Prior Work

Often, the problem of volumetric reconstruction of the mouse brain is reduced to finding the appropriate 2D registration algorithm for aligning consecutive slices, and the 3D volume is constructed by concatenating the transformations resulting from pairwise registrations (e.g., [2]). As argued in Sec. 3.2, good interslice alignment does not guarantee accurate reconstruction of the 3D brain shape. In [3, 4], such reconstruction is followed by non-rigid 3D registration to MR microscopy data. In [5], the reconstruction problem is formulated as a simultaneous system of 2D elastic registrations, for which an efficient solver is available. The input to this system is a rigid reconstruction that uses the principal axis transformation, which is driven by the shape of slice outlines. The rigid reconstruction method presented in this paper is intensity-driven and should provide a better initialization for elastic registration. In the Harvard High Resolution Brain Atlas [6], 3D reconstruction is aided by landmarks that are identified manually. Similarly, [7] uses surface registration and hand-drawn contours to warp MRI data to histology. In contrast, our algorithm is automatic, allowing high throughput.

## 3 Materials and Methods

### 3.1 Histology and Reference MRI Data

The brain from a 56 day old sacrificed C57BL/6J strain mouse was surgically removed, chemically frozen and sectioned in the coronal plane into 525 0.25 $\mu$m thick slices using a microtome. Sections were fixed, treated with Nissl counterstain that stains nucleic acids in neuronal somata and dendrites, and imaged, producing high-resolution color images with pixel size 0.95 $\mu$m$^2$. While histological images capture incredible anatomical detail (Fig. 1d), they also tend to suffer from artifacts, such as stretching, tearing and displacement of tissue that occurs during microtomy, as well as debris that appear during slice preparation and staining. These artifacts are illustrated in Fig. 1a,b.

In addition to the histological data, we use a 3D reference volume that was constructed by averaging a set of 30 *in vivo* MRI scans from 10 mice of the same strain as the specimen used to generate histology data. The reference volume has voxel size 12.9 $\mu$m$^3$ and is shown in Fig. 3.

### 3.2 Coarse-to-Fine Reconstruction Overview

Due to the large number of slices and high incidence of artifacts, the straightforward approach of aligning and warping each slice to its neighbors and concatenating the resulting transformations does not yield acceptable results. In particular, the accumulation of errors can result in the *z-shift effect*, where, though each slice is registered well to its neighbors, the overall 3D shape of the reconstruction is distorted, such that the imaginary grid lines parallel to the $z$-axis in the true anatomy become curves in the reconstruction. Another problem is the propagation of errors due to the presence of highly distorted sections

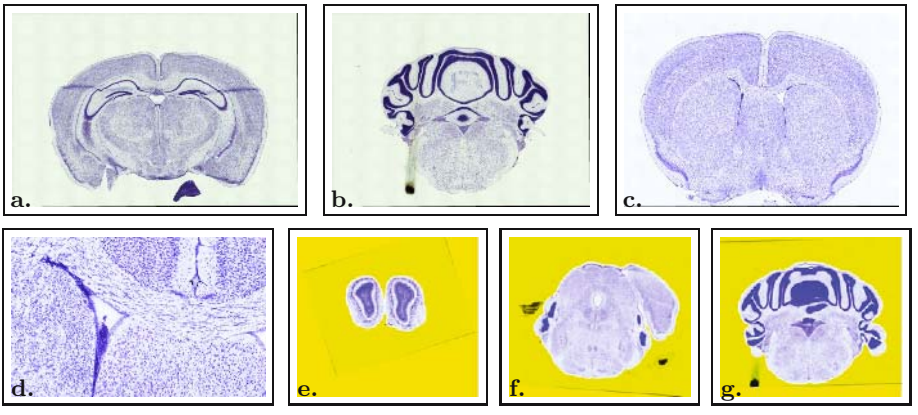



**Fig. 1. a-d:** Examples of Nissl-stained histological sections used in the reconstruction (**a:** a slice with tissue distortion; **b:** a slice with debris; **c:** an artifact-free slice; **d:** a subregion of slice **c** at full resolution). **e-f:** Examples of masks computed automatically for each slice. The region falling outside the mask is shaded yellow.

that are unlikely to register well to their neighbors and can cause gross discontinuities in the reconstructed anatomy. To address these issues, we developed a multi-stage method that combines segmentation, coarse-to-fine rigid reconstruction and deformable reconstruction to generate an atlas that is accurate both in terms of local anatomical continuity and global 3D brain shape.

We begin by giving an outline of our approach, with each step detailed in the subsequent sections. In the first stage, we use active contour segmentation to compute binary masks of the brain in each slice, so as to keep debris and dislodged tissue from contributing to the reconstruction. Next, we perform pairwise rigid registrations between neighboring slices to generate a locally accurate, fine-scale estimate of the reconstruction. By computing transforms not only between consecutive slices but between slices within a certain $z$-distance from each other, we reduce the $z$-shift and avoid negative effects of highly distorted slices. This initial 3D reconstruction is then registered to the reference MRI, producing a globally accurate, coarse-scale estimate of the reconstruction. The high-frequency component of the fine-scale estimate is then combined with the low-frequency component of the coarse-scale estimate to produce the final coarse-to-fine rigid reconstruction. This rigid reconstruction is then used to initialize a deformable registration, which produces a high quality 3D atlas.

### 3.3 Automatic Mask Computation

The background in histological slices can be highly inhomogeneous, containing debris, stains, and the edges of the glass plate. To keep these features from influencing slice alignment, we separate the brain from the background in each slice automatically, using a combination of active contour level set segmentation with

region competition [8] and mathematical morphology. The image force governing contour evolution is based on an estimate of object and background probabilities, which are estimated using thresholds. Active contour segmentation separates the brain from the more homogeneous background regions and island-like debris but it does not separate it from artifacts that are adjacent to it, such as the smear in Fig. 1c. We use erosion followed by dilation and taking the largest connected component to mask out such artifacts. Parameters of this masking approach may vary from slice to slice; however, in practice, a common set of parameters 'works' for over 90% of the slices; for the remaining slices, the researcher adjusts the parameters after examining the segmentation results. It may be possible to fully automate parameter selection in the future by using histogram analysis to set thresholds and by requiring masks computed on consecutive slices to have similar shape. Examples of masks are shown in Fig. 1e-g.

### 3.4 Fine-Scale Alignment

Our fine-scale alignment algorithm aims to minimize $z$-shift and the negative effects of badly distorted slices. This is accomplished by

1. Registering each histology slice not only to its nearest neighbors but also to neighbors located up to 5 slices away. The Insight Toolkit rigid registration pipeline [9] with the Mattes et al. mutual information metric [10] is used.
2. Constructing a weighted graph where the vertices represent the slices, the edges represent registrations between neighboring slices and the edge weights reflect the misregistration error, as detailed below.
3. Designating one slice as the reference and finding the shortest path from every vertex in the graph to the reference. These paths skip over those slices which register poorly to their neighbors, as Fig. 2 illustrates. For each slice, the chain of rigid transformations corresponding to the shortest path is concatenated, providing a rigid transformation from the slice to the reference.

The weight of the graph edge connecting vertices $i$ and $j$ is given by

$$W_{ij} = (1 + M_{ij})|i - j|(1 + \epsilon)^{|i-j|} \ , \tag{1}$$

where $M_{ij}$ is the value of the mutual information metric mapped to the range $[0, 1]$ (smaller values correspond to better registration), and $\epsilon$ is a positive constant that modulates between too little slice skipping, which results in propagation of registration error due to distorted slices, and too much slice skipping, which can disturb fine-scale alignment between neighbor slices. When $M_{ij}$ for all slice pairs are equal, no slices are skipped.

### 3.5 Coarse-Scale Alignment

To recover the gross shape of the mouse brain, we compute rigid registrations between histology slices and corresponding cross-sections of the reference MRI atlas. To compute the correspondences, we first use rigid 3D registration to

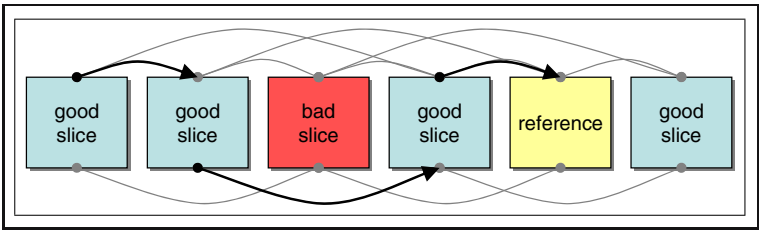



**Fig. 2.** An illustration of the graph-based reconstruction approach where each slice is registered to several neighbors and 'bad' slices (those that register poorly to neighbors) are skipped when aligning 'good' slices to the reference. Gray arcs indicate edges in the graph, i.e., transforms computed using rigid registration, and black arrows show the shortest path from the leftmost slice to the reference.

align the MRI volume to the histology volume reconstructed by the fine-scale algorithm; we then resample the MRI in the space of the histology volume, and extract 525 slices in the $z$-dimension. Both the 3D MRI-to-histology registration and the subsequent 2D histology-to-MRI registrations use the Mattes mutual information metric.

### 3.6 Coarse-Fine Recombination

The coarse and fine approaches described above produce 3D volumes that each have their own inadequacies. Fine-scale reconstruction aligns features well between neighboring slices, but does not eliminate $z$-shift. The coarse reconstruction attempts to recover the true shape of the brain, but it does so at the cost of local accuracy, as the registrations taking two neighboring histology slices into corresponding MRI slices are likely to disturb the alignment between the histology slices. In order to generate a single volume that fuses the attractive features of both reconstructions, we combine the high-frequency component of the fine-scale reconstruction with the low-frequency component of the coarse-scale reconstruction. This is achieved simply by smoothing across the $z$ dimension the parameters of the transforms that map histology slices into corresponding MRI slices. Gaussian smoothing with $\sigma = 100\mu$m is applied to each component of the rigid transforms between histology and MRI. This makes the mapping from the fine-scale histology reconstruction into the MRI volume smooth across the $z$ axis, maintaining the local correspondences established during slice-to-slice histology registration, while shifting the overall shape of the reconstructed volume to match that of the MRI brain. The choice of $\sigma$ was made empirically by estimating that the misregistration error of histology-to-MRI registration is approximately equal to four times the error in histology-to-histology registration.

### 3.7 Diffeomorphic Reconstruction

The deformable component of our reconstruction method will be presented elsewhere and is not the focus of the present paper. We summarize it here for the

sake of completeness. The reconstruction borrows the ideas from mesh fairing algorithms, where meshes are smoothed by moving each vertex to the average of its neighbors over a number of iterations. In our iterative method, each slice is deformed towards the average of its neighbors using a diffeomorphic inverse consistent algorithm [11]. This iterative deformation essentially smooths the shape of anatomical structures across the $z$ dimension. Akin to variable conductance diffusion, badly distorted slices, which are identified during rigid reconstruction, are assigned low weights in the iterative shape averaging process, preventing the distortions from propagating into neighboring slices, while allowing the good slices to influence the deformation of their badly distorted neighbors.

## 4 Results

Sagittal and horizontal cross-sections of the fine-scale, coarse-scale and coarse-to-fine rigid reconstructions are shown in Fig. 3. The thick blue outline indicates

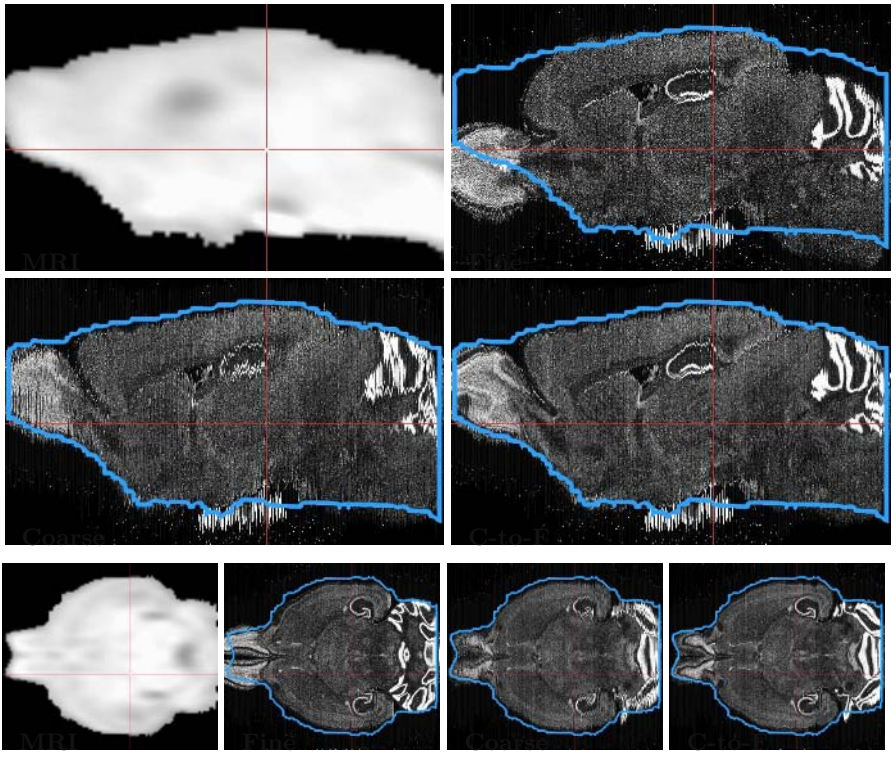


**Fig. 3.** MRI-based reference image and results of rigid histology reconstruction, including fine-scale reconstruction (slice-to-slice histology), fine-scale reconstruction (histology-to-MRI) and combined coarse-to-fine reconstruction. The outline of the brain surface in the MRI reference is overlaid as a blue curve on the reconstruction results.

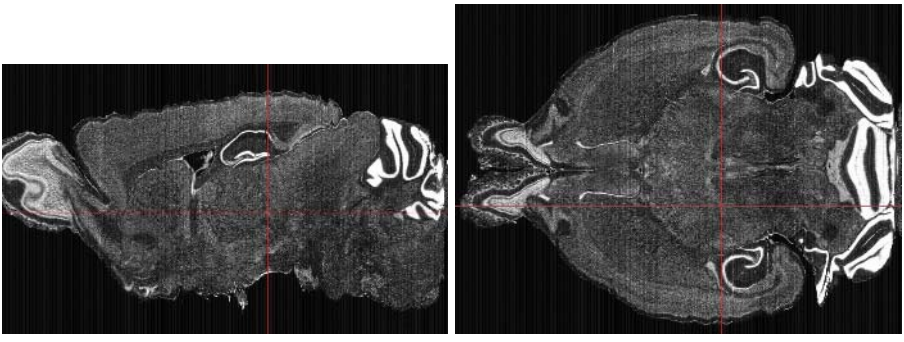


**Fig. 4.** Deformable reconstruction result

the boundary of the brain in the MRI reference. The $z$-shift occurring in fine-scale reconstruction is clearly evident, particularly in the sagittal cross-section. Poor local alignment of the coarse-scale rigid reconstruction is evident when comparing the cerebellum and other high-contrast structures between fine-scale and coarse-scale reconstructions. Visual inspection of the coarse-to-fine rigid reconstruction shows that it is superior to both fine-scale and coarse-scale reconstructions. The results of the deformable reconstruction are shown in Fig 4. Given the excellent initialization provided by the rigid method, the deformable method produces a reconstruction that is accurate in terms of both the overall 3D shape and the continuity of anatomical features across the $z$ dimension.

## 5 Discussion and Conclusions

A drawback of our method is the number of parameters that are chosen empirically. These include parameters for mask computation, the factors in computing the weights of the graph in fine-scale reconstruction and the width of the Gaussian filter used in the coarse-fine combination. It is difficult to set optimal values for these parameters because the ground truth against which the reconstruction result could be compared is not available. To a great extent, we rely on visual inspection to assess reconstruction quality but our future work will focus on generating quantitative assessments. One potential approach is to measure how well our reconstruction from coronal slices can match sagittal sections obtained from specimens from the same strain. This 2D/3D registration metric is especially useful because the ABA reference atlas itself is subsequently used as a reference to which slices from sagittal *in situ* hybridization are registered.

In conclusion, we have presented the details of a coarse-to-fine rigid 3D reconstruction technique for histological data. The approach combines the local accuracy of intensity-based registration between neighboring slices with the global accuracy of registration to a reference MRI dataset. Our results indicate that the combined coarse-to-fine approach is superior to either of its coarse and fine components taken on their own. Followed up with deformable diffeomorphic

reconstruction, our approach generates a high quality anatomical murine brain atlas that has been adopted by the Allen Institute for Brain Science as the reference for mapping gene expression.

## Acknowledgement

This work was made possible through funding support from the Allen Institute for Brain Science.

## References

1. Lydia Ng, Michael Hawrylycz, and David Haynor. Automated high-throughput registration for localizing 3D mouse brain gene expression using ITK. *Insight Jounral*, 1, 2005. Special Issue on ISC/NA-MIC/MICCAI Workshop on Open-Source Software.
2. S. Ourselin, A. Roche, G. Subsol, X. Pennec, and N. Ayache. Reconstructing a 3D structure from serial histological sections. *Image and Vision Computing*, 19(1):25–31, Jan 2001.
3. Allan MacKenzie-Graham, Erh-Fang Lee, Ivo D Dinov, Mihail Bota, David W Shattuck, Seth Ruffins, Heng Yuan, Fotios Konstantinidis, Alain Pitiot, Yi Ding, Guogang Hu, Russell E Jacobs, and Arthur W Toga. A multimodal, multidimensional atlas of the C57BL/6J mouse brain. *J Anat*, 204(2):93–102, Feb 2004.
4. S. Ourselin, E. Bardinet, D. Dormont, G. Malandain, A. Roche, N. Ayache, D. Tande, K. Parain, and J. Yelnik. Fusion of histological sections and mr images: Towards the construction of an atlas of the human basal ganglia. In *Medical Image Computing and Computer-Assisted Intervention, MICCAI*, pages 743–751, London, UK, 2001. Springer-Verlag.
5. Stefan Wirtz, Nils Papenberg, Bernd Fischer, and Oliver Schmitt. Robust and staining-invariant elastic registration of a series of images from histologic slices. In J. Michael Fitzpatrick and Joseph M. Reinhardt, editors, *Medical Imaging 2005: Image Processing*, volume 5747, pages 1256–1262. SPIE, 2005.
6. R.L. Sidman, B. Kosaras, B. Misra, and S. Senft. High resolution mouse brain atlas, 2005. http://www.hms.harvard.edu/research/brain.
7. M. A. Jacobs, J. P. Windham, H. Soltanian-Zadeh, D. J. Peck, and R. A. Knight. Registration and warping of magnetic resonance images to histological sections. *Med Phys*, 26(8):1568–78, Aug 1999.
8. Song Chun Zhu and Alan Yuille. Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(9):884–900, 1996.
9. L. Ibanez, W. Schroeder, L. Ng, and J. Cates. *The ITK Software Guide*. Kitware, Inc., 2003.
10. David Mattes, David R. Haynor, Hubert Vesselle, Thomas K. Lewellyn, and William Eubank. Nonrigid multimodality image registration. In Milan Sonka and Kenneth M. Hanson, editors, *Medical Imaging 2001: Image Processing*, volume 4322, pages 1609–1620. SPIE, 2001.
11. B. Avants, P.T. Schoenemann, and J.C. Gee. Lagrangian frame diffeomorphic image registration: Morphometric comparison of human and chimpanzee cortex. *Medical Image Analysis*, 2005. In Press.

# A Generalization of Free-Form Deformation Image Registration Within the ITK Finite Element Framework

Nicholas J. Tustison, Brian B. Avants, Tessa A. Sundaram,
Jeffrey T. Duda, and James C. Gee

Penn Image Computing and Science Laboratory
University of Pennsylvania
Philadelphia, PA 19104-6389

**Abstract.** Since the 1970's B-splines[1] have evolved to become the *de facto* standard for use in curve and surface representation. This resulted in a relatively recent proliferation of nonrigid image registration techniques based on B-splines. These techniques fall under the general Free-Form Deformation (FFD) approach in which the object to be registered is embedded within a B-spline object. The deformation of the B-spline object represents the transformation space of the registered object. In this paper, we describe the implementation of our finite element methodological (FEM) approach using B-splines. This registration framework subsumes essential components of currently popular FFD image registration algorithms, while providing a more principled and generalized control mechanism for nonrigid deformation. Our implementation constitutes an extension of the existing FEM library of the Insight Toolkit (ITK). We discuss the theoretical implications and provide experimental results of our proposed methodology.

## 1 Introduction

FFD approaches to modeling are characterized by embedding the object of interest within an encompassing geometric primitive. The resulting deformation of the object occurs via the transformation of the surrounding geometric space. Due, in large part, to their "local support" property and computational efficiency, FFD image registration based on B-splines has received substantial attention in the research literature. In this paper, we discuss our FEM-based B-spline approach to image registration, which is a generalization of current FFD B-spline image registration algorithms. In this sense, we are not proposing a new FFD algorithm which distinguishes itself from existing algorithms by one of the features discussed below, but rather we are claiming that current approaches can be subsumed within our proposed FEM image registration framework.

The constitutive framework of the general registration algorithm is best elucidated within the registration schema outlined in [2] which follows closely that of [3] and [4]. According to [2], the three major components comprising the general image registration algorithm are:

[1] Contrary to popular usage in the CAGD community, we follow de Boor's distinction discussed in [1] of using the term 'B-spline' to denote the shape function of limited support and not the B-spline object (*e.g* curve, surface, volume).

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 238–246, 2006.

- the *similarity metric*,
- the *optimization strategy*, and
- the *transformation model*.

While the transformation model addressed in this paper is B-spline-based, it is important, in distinguishing previous work, to discern the absence or presence of an explicit regularization component of the transformation. In addition, discussion of the first two components will prove helpful in differentiating previous work.

The explicit regularization term is best understood by interpreting the image registration problem in a Bayesian decision-theoretic context [5]. Potential solutions are viewed as posterior probabilities defined in terms of both the data (likelihoods) and the regularization (prior probabilities) components. The often-used variational formulation follows by employing an assumption of Gaussianity for the probability distribution functions. We write this general variational energy, $\Pi$, as

$$\Pi(I, J, T) = \int_{\Omega} (\Pi_{\sim}(I, J, T)(\mathbf{x}) + \Pi_R(T)(\mathbf{x}))\, \mathrm{d}\Omega, \tag{1}$$

where $I$ and $J$ are the fixed and moving images, respectively, and $T$ is the transformation which maps between $I$ and $J$. $\Pi_{\sim}$ is the similarity metric and $\Pi_R$ is the explicit regularization term. Some of the algorithms discussed below rely solely on the B-spline transform, $T$, to provide the desired continuity, at the exclusion of the $\Pi_R$ contribution.

## 2 Previous Work

Both Szeliski *et al.* [6] and Thévenaz *et al.* [7] use the Levenberg-Marquardt optimization strategy to formulate their respective B-spline image registration algorithms. The latter work was extended in [8] to include an explicit regularization term designed to minimize the weighted divergence and curl of the resulting deformation field. Unlike the aforementioned algorithms, which seek to minimize a mean squares image intensity similarity metric (MSQ), Rueckert *et al.* [9] and Mattes *et al.* [10] use normalized mutual information (NMI). The former method uses a gradient descent optimization strategy whereas the latter is optimized via a limited-memory, quasi-Newton minimization package (LBFGS). Also, the algorithm in [9] employs an explicit regularization term in the form of the well-known thin-plate energy to prevent topological folding of the deformation field for high-resolutions of the control point grid. This work was extended in [11] to include an isochoric regularization term, *i.e.* a regularization term which penalizes deviations of the Jacobian from unity. These algorithms are summarized in Table 1. By taking advantage of the modular nature of the ITK image registration framework [3], the basic elements of each of these algorithms can be reimplemented using ITK with relative ease.

While these FFD methods are geometrically-based, finite element methods are inherently physics-based. Physics-based approaches to image registration were pioneered by Broit and Bajcsy [12] in which image registration is modeled via the deformation of a physically-defined elastic object. Regularization occurs via the Cauchy-Navier operator derived from continuum mechanics, which is given by

$$L = \mu\nabla^2 + (\lambda + \mu)\nabla(\nabla\cdot), \tag{2}$$

**Table 1.** Summary of FFD B-Spline Registration Algorithms

| Method | Similarity Metric | Optimization Strategy | Explicit Regularization |
|---|---|---|---|
| Szeliski [6] | MSQ | Levenberg-Marquardt | no |
| Thévenaz [7] | MSQ | Levenberg-Marquardt | no |
| Sorzano [8] | MSQ | Levenberg-Marquardt | Divergence and Curl |
| Rueckert [9] | NMI | Gradient Descent | Thin-Plate |
| Mattes [10] | NMI | LBFGS | no |
| Rohlfing [11] | NMI | Gradient Descent | Isochoric |

where the Lamé constants, $\lambda$ and $\mu$, specify the material properties. This approach was later recast in variational form in [5] (see Equation (1)). This reformulation suggests the use of well-established finite element solution methods. Currently, this strategy is freely available for use and development as part of the ITK FEM software library. The implementation of B-spline finite elements within the ITK library allows for a generalized approach to B-spline image registration where the similarity metrics and regularization terms are easily tailored, in a modular fashion, to the problem of interest. This also allows for the essential components of the algorithms listed in Table 1 to be duplicated within our framework.

## 3 FEM Subsumption of FFD B-Spline Image Registration

The outline given of previous work provides the necessary context for understanding the subsumption of the major components of existing FFD methods into our FEM-based method.

### 3.1 Similarity Metric

Numerous similarity metrics are available for image registration in addition to the metrics used by the algorithms given in Table 1. Measures such as normalized cross correlation, Mattes' mutual information, and pattern intensity, as well as others, are all available within the ITK library. Instead of creating a new FFD algorithm for each metric, our FEM-based image registration framework allows for modular exchange of metrics dictated by the user.

### 3.2 Optimization Strategy

Although not explicit in the previous work listed, except for [6], it is assumed that finite difference approximations were used to estimate the value of the cost function and its derivatives for nonlinear optimization. Although justification is typically derived from a Taylor expansion perspective, finite differences can be thought of in terms of generalized FEM. From this vantage, finite difference solutions are simply FEM approximations based on localized, discontinuous shape functions. Although such methods are easy to implement, theoretically speaking, they are less accurate than corresponding FEM approaches [13].

### 3.3 Explicit Regularization

Of the six algorithms surveyed in Table 1, three use no explicit regularization. These algorithms depend on the inherent continuity of higher order B-splines to satisfy smoothness constraints. Of the other three, the work in [9] uses only an extremely tiny contribution of the thin-plate energy regularization. The weighting value, $\lambda = 0.01$, was determined empirically and is only used when the control point density is sufficiently large (although the determination of "sufficiently large" is left unexplained). In [11] it was noted that the thin-plate energy has no inherent volume-preserving capabilities which motivated the introduction of a Jacobian-based regularization term. Unfortunately, there are scenarios in which this assumption is invalid, *e.g.* intersubject brain registration. In [8] the capability of regularizing in terms of both the divergence and curl of the displacement field is proposed. For many of the results given in the paper, however, there was no contribution from the curl regularization term whereas the divergence term had a weighting of 4 (assumed to be set empirically). In contrast to the algorithm-specific regularization of the listed FFD methods, FEM-based image registration allows for simple, modular exchange of regularization types (*e.g.* membrane or thin-plate regularization).

## 4 FEM Image Registration Using B-Splines

### 4.1 Shape Functions and Nodal Placement

At its conceptual level, FEM solutions discretize the continuum of the problem domain into *elements*. Each of these elements has a governing set of both nodes (or control

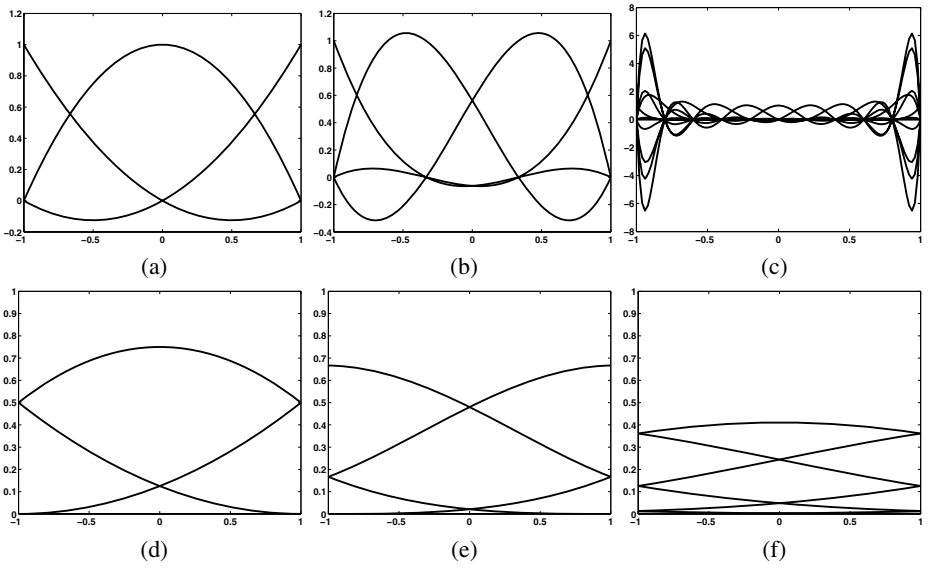



**Fig. 1.** Shape functions: (a) $C^1$ Lagrangian, (b) $C^2$ Lagrangian, (c) $C^9$ Lagrangian, (d) $C^1$ B-spline, (e) $C^2$ B-spline, and (f) $C^9$ B-spline

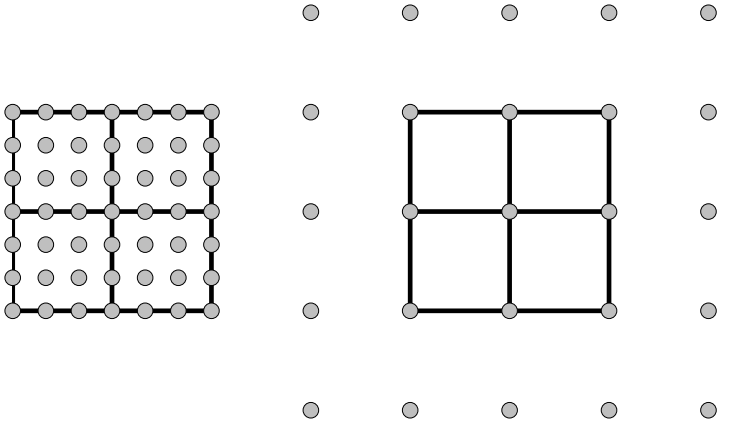


**Fig. 2.** Placement of nodes for Lagrangian (left) versus B-spline elements (right)

points) and shape functions. In calculating the value of the FEM solution at a specific point within a given element, the shape functions are used to average, in a weighted and smooth fashion, the contribution of the corresponding nodal values. Finite elements using B-splines, and a comparison with more traditional elements, have been investigated previously [14] in a general FEM context. We compare B-spline elements with Lagrangian elements, due to their historical ubiquity in the FEM literature [13], in order to provide a point of reference in discussing B-spline elements.

For simplicity in comparing shape functions, we plot the univariate shape functions of both the Lagrangian and B-spline varieties. Multivariate extensions are simply the tensor product of the respective univariate set. These shape functions, parameterized over the interval $[-1, 1]$, are illustrated in Figure 1. Note that, when specified, the FFD B-spline image registration algorithms surveyed use cubic B-splines (see Figure 1(e)).

The Lagrangian shape functions are interpolatory in that the function values are identical with the nodal values at the nodal locations. This is, in general, not the case with B-spline elements (the linear case is the sole exception in which Lagrangian and B-spline shape functions are identical). Also, notice that the B-spline shape functions are strictly nonnegative over the interval in contrast to their Lagrangian counterparts. The Lagrangian shape functions also exhibit undulations meaning that the nodal influence in calculating the elemental function value does not necessarily decrease monotonically with distance as it does with B-splines. This undulatory phenomenon is exacerbated in higher order shape functions, *e.g.* Figure 1(c).

In addition to the distinct shape functions, nodal location differs between Lagrangian and B-spline elements. For a given mesh size, B-spline elements span a subspace of the Lagrangian elements so the theoretical accuracy will necessarily be less [14]. However, where computational costs prohibit fine resolution of Lagrangian element meshes (as in the case with certain image registration problems), B-spline elements provide a less costly substitute while guaranteeing a specified order of continuity. Shown in Figure 2 are two diagrammatic representations of the nodal placement for four $C^2$ Lagrangian elements (left) and four $C^2$ B-spline elements (right). Note that neighboring Lagrangian

elements only share bordering nodes whereas neighboring $C^2$ B-spline elements share 12 common nodes.

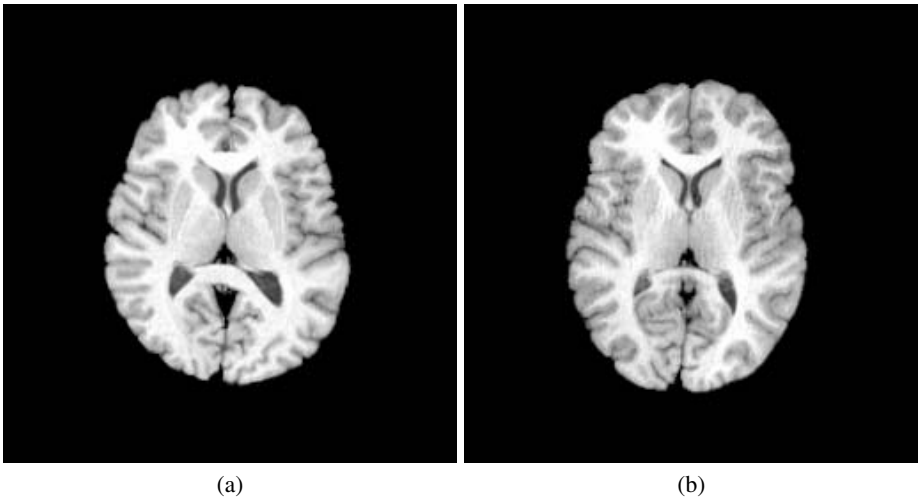
### 4.2 Multiresolution Strategy

The current ITK FEM registration methodology allows for a mulitresolution approach in which more global deformation can be inferred using a low resolution grid which can then be propagated to higher levels for more localized deformation. We generalize the hierarchical approach for doubling the resolution of cubic B-spline surfaces discussed in [15] to bivariate and trivariate B-spline objects of arbitrary order. This allows for propagating the FEM B-spline registration solution at a lower level to the next higher level.

## 5 Experimental Results

We experimentally demonstrate results from our image registration framework using two 2-D brain slice images (size $= 256 \times 256$) from two different subjects. These images are shown in Figures 3(a) and 3(b) which comprise the moving and fixed images, respectively. Preprocessing of the images included histogram matching. The mean square image intensity similarity metric was used in each of the examples as was a single level resolution. Two different grid sizes were used ($64 \times 64$ and $128 \times 128$ elements). These grids were composed of quadratic ($C^1$) B-spline elements using both membrane and linear elastic regularization models [13].

The image registration results are illustrated in Figure 4. The top row consists of the results obtained from the membrane regularization model whereas the bottom row



(a) (b)

**Fig. 3.** 2-D Brain slice images used to demonstrate the FEM-based B-spline image registration methodology. (a) Original moving image and (b) fixed image.

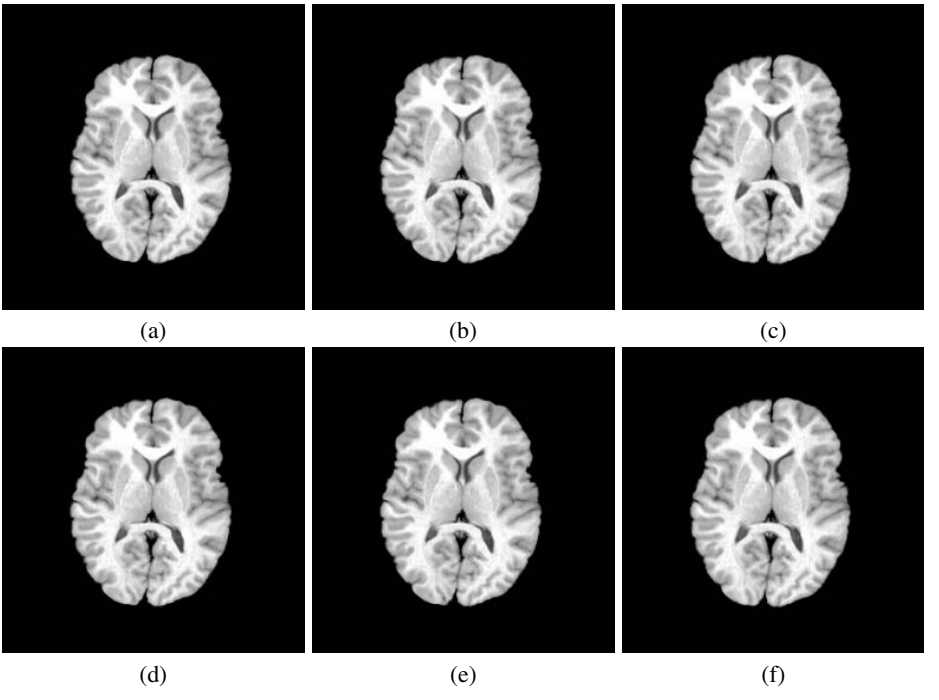



**Fig. 4.** Registered images using different B-spline elements and Lamé constants: (a) membrane element, $\mu = 25000$, $\lambda = 0$, $64 \times 64$ nodal mesh; (b) membrane element, $\mu = 25000$, $\lambda = 0$, $128 \times 128$ nodal mesh; (c) membrane element, $\mu = 50000$, $\lambda = 0$, $128 \times 128$ nodal mesh; (d) linear elastic element, $\mu = 25000$, $\lambda = 0$, $64 \times 64$ nodal mesh; (e) linear elastic element, $\mu = 25000$, $\lambda = 0$, $128 \times 128$ nodal mesh; (f) linear elastic element, $\mu = 50000$, $\lambda = 0$, $128 \times 128$ nodal mesh;

contains the results from the linear elastic model. The first column of image results in Figure 4 were derived from grids of size $64 \times 64$ elements whereas the last two columns of images were derived from grids of size $128 \times 128$ elements. Experimentation led to the selection of the Lamé constants used for the membrane results. Lower values for $\mu$ in the membrane model caused folding of the deformation field which dictated our selection of the given parameters. We employed the same $\mu$ and $\lambda$ values used in the membrane model for the linear elastic model for comparative purposes.

To assess the quantitative difference between the resulting images, we list various measures in Table 2. These measures include the mean squares intensity difference between the resulting image and the fixed image as well as the minimum and maximum Jacobian of the resulting transformation. These numbers seem to accord with intuition. The results derived from a higher resolution grid have a lower mean squares error value. Also, we see that the membrane regularization produces deformation fields with extreme Jacobian values closer to unity and lower mean squares error values than the linear elastic regularization.

**Table 2.** Quantitative measures for the results in Figure 4

| | Membrane $64 \times 64$ $\lambda = 0$ $\mu = 25000$ **(a)** | Membrane $128 \times 128$ $\lambda = 0$ $\mu = 25000$ **(b)** | Membrane $128 \times 128$ $\lambda = 0$ $\mu = 50000$ **(c)** | Linear El. $64 \times 64$ $\lambda = 0$ $\mu = 25000$ **(d)** | Linear El. $128 \times 128$ $\lambda = 0$ $\mu = 25000$ **(e)** | Linear El. $128 \times 128$ $\lambda = 0$ $\mu = 50000$ **(f)** |
|---|---|---|---|---|---|---|
| Mean squares | 237.7 | 190.0 | 193.0 | 221.7 | 175.1 | 206.1 |
| Min. Jacobian | 0.72 | 0.52 | 0.42 | 0.65 | 0.28 | 0.53 |
| Max. Jacobian | 1.36 | 1.73 | 1.76 | 1.48 | 1.76 | 1.61 |

## 6 Conclusion

We provided a discussion of our open-source ITK finite element library for nonrigid image registration based on B-splines and how it is related to current FFD B-spline image registration approaches. Experimental results were also illustrated using components unique to the FEM framework. Although specific FFD implementations are perhaps easier to implement, the FEM approach provides for more principled control over the nonrigid deformation. While we find the discussion presented in this paper meritorious, we recognize that the scope of our experimental results was minimal and that other areas of investigation would certainly be welcomed by the research community. Such areas of investigation would include, but not be limited to, the consequences of using other regularization models (*e.g.* thin plate splines), the affects of the multiresolution strategy, differences in the results produced by different orders of B-splines, or the affects of different similarity metrics. Such comprehensive discussion will certainly be included in future work.

## References

1. de Boor, C.: B-spline basics. In: Fundamental Developments of Computer-Aided Geometric Modeling. American Press (1993) 27–49
2. Crum, W.R., Hartkens, T., Hill, D.L.G.: Non-rigid image registration: theory and practice. The British Journal of Radiology **77** (2004) 140–153
3. Ibanez, L., Ng, L., Gee, J.C., Aylward, S.: Registration patterns: The generic framework for image registration of the Insight Toolkit. In: IEEE International Symposium on Biomedical Imaging. (2002) 345–348
4. Brown, L.G.: A survey of image registration techniques. ACM Computing Surveys **24** (1992) 325–376
5. Gee, J.C., Bajcsy, R.K.: Elastic Matching: Continuum Mechanical and Probabilistic Analysis. In: Brain Warping. Academic Press (1999) 183–198
6. Szeliski, R., Coughlan, J.: Spline-based image registration. International Journal of Computer Vision **22** (1997) 199–218
7. Thevénaz, P., Ruttimann, U.E., Unser, M.: A pyramid approach to subpixel registration based on intensity. IEEE Transactions on Image Processing **7** (1998) 27–41

8. Sorzano, C.O.S., Thevénaz, P., Unser, M.: Elastic registration of biological images using vector-spline regularization. IEEE Transactions on Biomedical Engineering **52** (2005) 652–663
9. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: Application to breast MR images. IEEE Transactions on Medical Imaging **18** (1999) 712–721
10. Mattes, D., Haynor, D.R., Vesselle, H., Lewellen, T.K., Eubank, W.: PET-CT image registration in the chest using free-form deformations. IEEE Transactions on Medical Imaging **22** (2003) 120–128
11. Rohlfing, T., Jr., C.R.M., Bluemke, D.A., Jacobs, M.A.: Volume-preserving nonrigid registration of mr breast images using free-form deformation with an incompressibility constraint. IEEE Transactions on Medical Imaging **22** (2003) 730–741
12. Bajcsy, R., Broit, C.: Matching of deformed images. In: International Conference on Pattern Recognition. (1982)
13. Zienkiewicz, O.C., Taylor, R.L.: The Finite Element Method, Volume 1: The Basis. $5^{th}$ edn. Butterworth Heinemann (2000)
14. Sabin, M.A.: Spline Finite Elements. PhD thesis, Leeds University (1997)
15. Lee, S., Wolberg, G., Shin, S.Y.: Scattered data interpolation with multilevel b-splines. IEEE Transactions on Visualization and Computer Graphics **3** (1997) 228–244

# A Novel Projection Based Approach for Medical Image Registration

Ali Khamene[1], Razvan Chisu[2], Wolfgang Wein[2],
Nassir Navab[2], and Frank Sauer[1]

[1] Imaging and Visualization Dept., Siemens Corporate Research,
755 College Road East, Princeton NJ 08540, USA
{ali.khamene, sauer.frank}@siemens.com
[2] Computer Aided Medical Procedures (CAMP) Group, TU Munich
Boltzmannstr. 3, 85748 Garching, Germany
{chisu, wein, navab}@cs.tum.edu

**Abstract.** In this paper, we propose a computationally efficient method for medical image registration. The centerpiece of the approach is to reduce the dimensions of each image via a projection operation. The two sequences of projection images corresponding to each image are used for estimating the registration parameters. Depending upon how the projection geometry is set up, the lower dimension registration problem can be parameterized and solved for a subset of parameters from the original problem. Computation of similarity metrics on the lower dimension projection images is significantly less complex than on the original volumetric images. Furthermore, depending on the type of projection operator used, one can achieve a better signal to noise ratio for the projection images than the original images. In order to further accelerate the process, we use Graphic Processing Units (GPUs) for generating projections of the volumetric data. We also perform the similarity computation on the graphics board, using a GPU with a programmable rendering pipeline. By doing that, we avoid transferring a large amount of data from graphics memory to system memory for computation. Furthermore, the performance of the more complex algorithms exploiting the graphics processor's capabilities is greatly improved. We evaluate the performance and the speed of the proposed projection based registration approach using various similarity measures and benchmark them against an SSE-accelerated CPU based implementation.

## 1 Introduction

Automatic image registration is nowadays an essential component in medical imaging systems. The basic goal of intensity based image registration techniques is to align anatomical structures in different modalities. This is done through an optimization process, which assesses image similarity and iteratively changes the transformation of one image with respect to the other, until an optimal alignment is found [1]. Computation speed is a critical issue and dictate applicability

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 247–256, 2006.

of the technology in practice. On the other hand feature based methods are computationally more efficient, however, they are notoriously dependant on the quality of the extracted features from the images [2].

In intensity based registration, volumes can directly be aligned by computing a volumetric similarity measure accessing the voxel intensities at each iteration. Since the amount of computations performed at each iteration is high, the overall registration process is very slow. In the cases, where Mutual Information (MI) is used, sparse sampling of volume intensity could reduce the computational complexity while compromising the accuracy [3, 4]. In [5], authors propose a projection based method for 2D-2D image registration. In this method, the projections along the two axes of the image are computed. Horizontal and vertical components of the shift is then computed using one-dimensional cross correlation based estimator. They show that the method is robust in the presence of temporal and spatial noise and computationally efficient compared to the 2D correlation based shift estimator. In [6], authors propose to formulate 3D-3D registration cost function as the summation of three 2D-3D optimization cost functions. The optimization is then done concurrently on the sum of the cost functions, which are identically parameterized. Furthermore, images are preprocessed to extract a binary segmentation. Projection images from the binary segmentation are used for computing the similarity measures.

Our proposed approach is a combination of the methods in [5] and [6]. We compute the projection images from two volumes and setup a cost function to register these images within a space, which is a subset of the space of the original rigid registration transformations. We perform these registrations successively for various projection geometries in order to estimate all the registration parameters of the original problem. We further optimize the performance of projection computation and 2D-2D registration similarity computation by using GPUs. We perform a validation study comparing the accuracy and the speed of the proposed method with a traditional volumetric 3D-3D MI-based approach.

## 2 Method

### 2.1 Proposed Registration Method

Conventional volumetric rigid registration algorithms, optimize 6 degrees of freedom (DOF) transformation parameters in order to maximize the volumetric similarity measure.

$$\widetilde{\mathbf{T}^6} = \arg \max_{\mathbf{T}^6} \mathcal{S}^3 \left( I_f, \mathcal{T}^3_{\mathbf{T}^6}(I_m) \right). \tag{1}$$

where $\mathbf{T^6}$ is a six DOF homogenous transformation matrix, $\mathcal{T}^3$ is the six DOF mapping operator, $\mathcal{S}^3$ estimates the similarity metric between two volumes, and $I_f$ and $I_m$ are the fixed and moving volumetric data, respectively. Let us define an orthographic projection operator $\mathcal{P}$, which projects the volume points onto an image plane using a projection matrix $\mathbf{\Pi}$. Without loss of generality, assuming

that the volume world coordinates are located at the center of the fixed volume, any plane projection can be written as:

$$\mathbf{\Pi_P} = \mathbf{\Pi}_0\mathbf{P}, \tag{2}$$

where $\mathbf{\Pi}_0$ is a trivial $3 \times 4$ orthographic projection matrix with the principle axis along $z$ and $\mathbf{P}$ is a $4 \times 4$ homogenous transformation matrix encoding the principle axis of the orthographic projection matrix $\mathbf{\Pi_P}$. Since we only consider orthographic projections and assume that the center of the volume maps to the center of the plane of projection, the translation part of the matrix $\mathbf{P}$ is zero. Once we compute the projection images using operator $\mathcal{P}_\mathbf{P}$, which uses the projection matrix $\mathbf{\Pi_P}$, we can re-formulate the registration optimization as follows:

$$\widetilde{\mathbf{T}^3_\mathbf{P}} = \arg \max_{\mathbf{T}^3_\mathbf{P}} \mathcal{S}^2 \left( \mathcal{P}_\mathbf{P}(I_f), \mathcal{T}^2_{\mathbf{T}^3_\mathbf{P}}(\mathcal{P}_\mathbf{P}(I_m)) \right), \tag{3}$$

where $\mathbf{T^3_P}$ is a three DOF homogenous transformation matrix defined in the plane of projection specified by $\mathbf{P}$, $\mathcal{T}^3$ is a three DOF mapping operator, $\mathcal{S}^2$ computes the similarity metric between the 2D images. There exists an explicit relationship between the reduced dimension homogenous transformation matrix $\mathbf{T}^3_\mathbf{P}$ and the original six DOF transformation matrix $\mathbf{T^6}$. Let us formulate $\mathbf{T}^3_\mathbf{P}$ as follows:

$$\mathbf{T}^3_\mathbf{P} = \begin{bmatrix} cos(\theta) & -sin(\theta) & t_x \\ sin(\theta) & cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

where $\theta$ is in-plane rotation and $t_x$ and $t_y$ are in-plane translations. It can easily be shown that:

$$\mathbf{T}^6 = \mathbf{P}^{-1}\mathbf{\Pi}_0^\top \mathbf{T}^3_\mathbf{P}\mathbf{\Pi}_0\mathbf{P} \tag{5}$$

where $\top$ depicts the transpose operation. From equation 5, it is apparent that only part of the space represented by $\mathbf{T}^6$ is covered by the in-plane transformation from $\mathbf{T}^3$. And that depends on the projection transformation in $\mathbf{P}$. In-plane translations from $\mathbf{T}^3$ are translations along the first two axes of the transformed coordinate system by $\mathbf{P}$ and the in-plane rotation is the rotation about the third axis. In order to cover the whole space of 3D rigid transformation, one easy solution is to consider three projections, as follows:

$$\begin{aligned} \mathbf{\Pi_{P_0}} &= \mathbf{\Pi}_0\mathbf{P}_0 = \mathbf{\Pi}_0 \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{0} \\ & \mathbf{0}^\top & & 1 \end{bmatrix} \\ \mathbf{\Pi_{P_1}} &= \mathbf{\Pi}_0\mathbf{P}_1 = \mathbf{\Pi}_0 \begin{bmatrix} \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{r}_1 & \mathbf{0} \\ & \mathbf{0}^\top & & 1 \end{bmatrix} \\ \mathbf{\Pi_{P_2}} &= \mathbf{\Pi}_0\mathbf{P}_2 = \mathbf{\Pi}_0 \begin{bmatrix} \mathbf{r}_3 & \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{0} \\ & \mathbf{0}^\top & & 1 \end{bmatrix} \end{aligned} \tag{6}$$

where $\mathbf{0}$ is vector of zero and $\mathbf{r}_i$ for $i \in [1\ 3]$ are columns of the rotation matrix embedded in the corresponding homogenous transformation. Finally the successive optimization approach is outlined in Algorithm 1.

precompute the projections of the fixed volume using $\mathbf{\Pi}_{\mathbf{P}_0}$ to $\mathbf{\Pi}_{\mathbf{P}_2}$
registration transformation is set to identity
**while** *there is a significant incremental change in registration parameters* **do**
    successively choose a projection matrix from $\mathbf{\Pi}_{\mathbf{P}_0}$ to $\mathbf{\Pi}_{\mathbf{P}_2}$
    compute the projection of the transformed moving volume
    perform registration between the corresponding projection images as in equation 3
    use the equation 5 to update the registration transformation
**end**

**Algorithm 1:** The proposed registration approach

Regardless of how computationally expensive a similarity measure estimation is, a volumetric registration method requires $k * N^3$ operations for an optimization loop with $k$ iterations, where $N$ is the number of voxels. For identical number of iterations, assuming two rounds of registrations per projection, the number of operations for the proposed method is $2 * N^3 + k * N^2$. Furthermore, the projection operation in this case can be performed using a graphics processing unit, as it is explained in the following section.

### 2.2 Implementation

The computation performance offered by today's video boards by far surpasses that of currently available CPUs - while a Pentium 4 3GHz CPU can theoretically reach 6 GFLOPS [1], synthetic benchmarks have shown the NVIDIA Geforce 6800 Ultra GPU to reach 40 GFLOPS. This fact, together with the inherently parallel architecture of graphics processors, has made the approach of GPU programming highly attractive for accelerating algorithms in different domains. For instance, robot motion planning [7], flow visualization [8], segmentation [9], and solving sets of algebraic equations [10] are proposed to be implemented using GPUs. In the domain of medical image registration, GPUs have been mainly used to speed up the generation of Digitally Reconstructed Radiographs (DRRs) using hardware-accelerated volume rendering techniques. Our aim is now to implement the proposed algorithm including the computation of image similarity using a GPU.

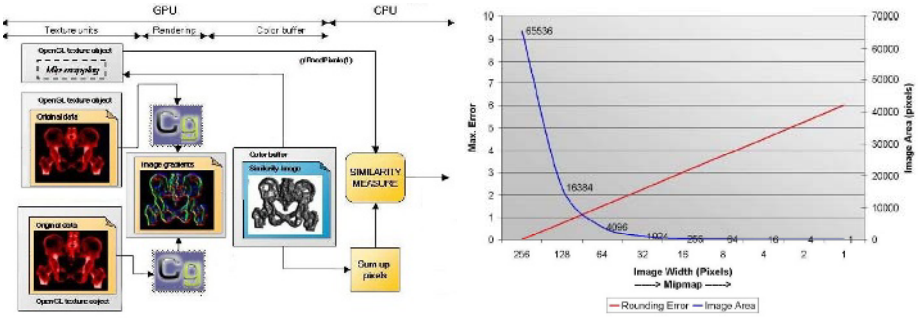
Graphics processors are highly specialized for a single application - handling three-dimensional geometric data, project it onto the two-dimensional computer display and apply a range of visual effects, like texturing, lighting and shading. There are two major steps in the rendering process. In the vertex shading step vertex-specific data like normals or texture coordinates are manipulated. Whereas, in the fragment shading step pixel-specific data like texturing or blending are processed. Both these steps can be customized by loading a so-called *shading-program* or *shader* onto the GPU. A different program is needed for each of these two steps. In the process of computing similarity measures, the major part, i.e., examining the image intensities, has to be done for each

[1] Billions of floating-point operations per second.

of the image pixels. This part is implemented as a fragment shader. However, the computation of the texture coordinate increments that have to be added to the current coordinate, in order to examine the pixel's neighborhood, can also be implemented within the vertex shading unit. Aside from these, all the other processes taking place in the rendering pipeline (e.g., rasterization) must be regarded as computational overhead. Therefore, simple similarity measures perform rather poorly in comparison to a classical CPU implementation, while for the more complex measures, i.e., the ones requiring a fragment shader with a high number of operations, the pipeline's overhead becomes less significant. Our results show that, depending on the measure's complexity, the GPU-based computation can be from three times slower to ten times faster than a non-accelerated CPU-based implementation.

The specialized nature of the GPU environment imposes several technical restrictions that do not apply to classic CPU programming. For example, the pipelined structure of the GPU implies that data can flow only in one direction, which means that the respective programs cannot modify their input values, nor can they read from the output buffers. This means that information-theoretic similarity measures like MI cannot be implemented. Furthermore, the pixels rendered onto the screen are usually represented as positive 8 bit RGBA color values. Thus in each rendering pass, up to four values can be computed for each pair of pixels in the images to be registered. Similarity measures requiring more values accordingly have to be split up into multiple rendering passes. Finally, with respect to the precision of the computed color data, 16 and 32 bit floating-point color buffers are also supported by the more modern video boards. However, these color buffers are always off-screen rendering buffers that cannot be directly displayed, and are both more difficult to use and impose a drop in performance. However, in the results section, we show that using a low-accuracy 8 bit color buffer can result in a registration just as accurate as using 32 bit floating point data on the CPU. Because 8 bit color intensities can take only positive values, any negative number automatically gets clamped to zero. As some of the measures we compute, and also the image gradients, yield negative values, we decided to use two color channels to store the value by employing the following scheme: $R = x, G = -x \Rightarrow x = R - G$.

The similarity of two images must be provided as a scalar value, e.g. the sum or average of all pixels, in order to represent a cost function to be optimized. However, summing up the color values in the frame buffer or within a texture is not a typical graphics application. In order to avoid the slow copying of the computed image from the GPU to the main RAM for averaging on the CPU, we use the mipmapping capability of the video card. Mipmaps represent successively smaller versions of an initial texture, having half the width and height of the previous mipmap. They are obtained by taking four neighbored pixels from the input image and averaging them into one pixel in the output image. The main problem of mipmaps arises from the fact that the repeated summation and averaging of 8 bit data (e.g., 8 times for a $256 \times 256$ image) is followed by a truncation/rounding to 8 bits. These rounding errors can potentially affect the




(a) (b)

**Fig. 1.** (a)Similarity measure computation process on GPU, and (b) depicts the rounding error and image size with respect to mipmapping level

registration accuracy. To work around this problem, we use a hybrid averaging approach. Instead of using the last mipmap consisting in a single pixel, we generate the 4th mipmap, which is $16 \times 16$ pixels in the case of $256 \times 256$ images, and then copy this very small image to the main memory for CPU averaging. By doing this, we significantly reduce the amount of data that has to be transferred from video to system memory without introducing more rounding errors than necessary, resulting in a good compromise between speed and accuracy.

We implemented ten similarity measures of different complexity: Sum of Squared Differences (SSD)[11], Sum of Absolute Differences (SAD), Correlation Coefficient (CC)[11], Ratio of Image Uniformity (RIU)[12], Pattern Intensity (PI)[13], Gradient Correlation (GC), Gradient Difference (GD)[14], Gradient Proximity (GP), Sum of Local Normalized Correlation (SLNC)[15] and Variance-Weighted Sum of Local Normalized Correlation (VWC)[15]. In order to compute the projection images, we used a 3D texture based volume rendering technique. Both intensity accumulation or maximum intensity projection (MIP) techniques can be used for generating the projection images. Vector valued projection images can also be derived from the volumes. These images provides much richer information regarding the volume at each specific projection. The proposed method can be used along with all the various volume projection options.

## 3 Results

We validated our methods by performing intensity-based registration on two data sets. The first comprised two CT scans of a patient's pelvis, before and after treatment, respectively. The second data was a CT scan of a megavolt cone-beam CT volume of a skull phantom. The standard registration approach, fully sampling the moving volume using the MI similarity metric, yielded very stable results and was thus used as ground truth. A registration run took several minutes using this method. A cube with $10cm$ width placed at the volume center

**Table 1.** Similarity Measure computation times

| Measure | CPU | CPU, SSE | GPU, mipmaps | GPU, CPU | GPU, hybrid |
|---|---|---|---|---|---|
| SAD | 0.17ms | 0.05ms | 0.54ms | 3.52ms | 0.60ms |
| SSD | 0.17ms | 0.05ms | 0.54ms | 3.54ms | 0.59ms |
| RIU | 1.20ms | 0.32ms | 0.55ms | 3.56ms | 0.62ms |
| NCC | 0.87ms | 0.17ms | 0.53ms | 3.56ms | 0.59ms |
| GC | 2.46ms | 0.55ms | 1.15ms | 7.12ms | 1.29ms |
| GD | 2.86ms | 0.70ms | 0.73ms | 3.71ms | 0.76ms |
| GP | 1.82ms | 0.45ms | 0.72ms | 3.70ms | 0.75ms |
| PI | 14.83ms | 7.93ms | 1.21ms | 4.20ms | 1.27ms |
| LNC | 18.32ms | 6.69ms | 3.14ms | 6.11ms | 3.17ms |
| VWC | 18.35ms | 6.59ms | 1.78ms | 4.77ms | 1.81ms |

**Table 2.** Registration results for pelvis CT

| | $GPU_{mm}1$ | $GPU_{mm}2$ | $GPU_{mm}3$ | $GPU_{CPU}1$ | $GPU_{CPU}2$ | $GPU_{CPU}3$ |
|---|---|---|---|---|---|---|
| Measure | GD | SAD | PI | PI | GD | RIU |
| TRE [mm] | 1.57 | 2.29 | 2.86 | 0.95 | 0.98 | 1.02 |
| $\sigma(TRE)$ | 0.60 | 0.83 | 4.3 | 0.22 | 0.31 | 0.21 |
| t [s] | 2.4 | 2.1 | 2.8 | 7.04 | 5.92 | 5.26 |

| | $GPU_{hyb}2$ | $GPU_{hyb}2$ | $GPU_{hyb}3$ | $CPU_{SSE}1$ | $CPU_{SSE}2$ | $CPU_{SSE}3$ |
|---|---|---|---|---|---|---|
| Measure | PI | GD | RIU | GD | PI | SAD |
| TRE [mm] | 0.9 | 0.98 | 1.04 | 0.82 | 0.98 | 1.01 |
| $\sigma(TRE)$ | 0.23 | 0.31 | 0.23 | 0.25 | 0.30 | 0.28 |
| t [s] | 4.00 | 3.5 | 2.9 | 6.76 | 13.02 | 6.57 |

was used for computing the Target Registration Error (TRE). Table 1 lists the computation times for computing the respective similarity measures for $256 \times 256$ pixel 8 bit grayscale images on the CPU with/without SSE-II acceleration, and on the GPU with mipmap and/or CPU-based averaging. These benchmarks were executed on an Intel Pentium4 2.4 Ghz system with an NVidia GeForce 6800GT graphics accelerator.

For validating our implementation, we used two pairs of volumetric data sets. We will present the three most accurate measures for four implementations: GPU-based using mipmaps for averaging ($GPU_{mm}$), GPU-based using the CPU for averaging ($GPU_{CPU}$), GPU-based using a combination of mipmaps and the CPU for averaging ($GPU_{hyb}$) and CPU-based with SSE ($CPU_{SSE}$) acceleration. First set of experiments are done on two CT scans of the same patient's pelvis taken at different points of time. The image intensity and contrast differed slightly, with some features being visible in only one of the volumes.

For the second run we used scans of a skull acquired with different types of CT scanners. The fixed volume was acquired using a regular CT scanner, while a megavolt cone-beam CT was used for the moving volume. Thus, the

**Table 3.** Registration results for skull CT / MV-CT

| | $GPU_{mm}1$ | $GPU_{mm}2$ | $GPU_{mm}3$ | $GPU_{CPU}1$ | $GPU_{CPU}2$ | $GPU_{CPU}3$ |
|---|---|---|---|---|---|---|
| Measure | SAD | SSD | RIU | SSD | NCC | SAD |
| TRE [mm] | 2.32 | 3.27 | 5.65 | 1.57 | 1.65 | 1.67 |
| $\sigma(TRE)$ | 0.71 | 1.94 | 3.58 | 0.09 | 0.18 | 0.04 |
| t [s] | 2.34 | 2.27 | 2.59 | 5.35 | 5.38 | 5.29 |

| | $GPU_{hyb}2$ | $GPU_{hyb}2$ | $GPU_{hyb}3$ | $CPU_{SSE}1$ | $CPU_{SSE}2$ | $CPU_{SSE}3$ |
|---|---|---|---|---|---|---|
| Measure | SSD | NCC | SAD | RIU | NCC | SSD |
| TRE [mm] | 1.68 | 1.85 | 1.86 | 1.37 | 1.52 | 1.54 |
| $\sigma(TRE)$ | 0.35 | 0.50 | 0.17 | 0.23 | 0.10 | 0.09 |
| t [s] | 2.99 | 2.98 | 2.99 | 7.19 | 6.21 | 6.49 |

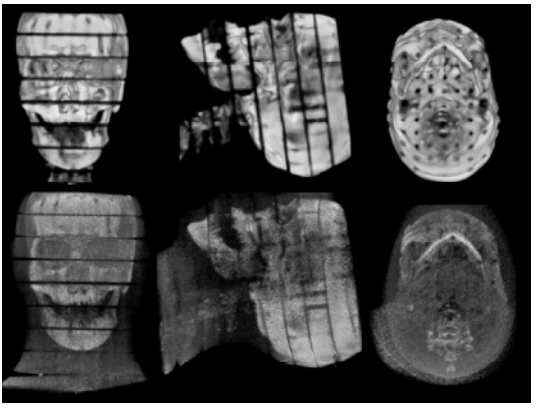


**Fig. 2.** First row depicts the projection images from a CT volume, second row shows the corresponding images from Mega-Voltage CT volume

two volumes differed quite strongly with respect to brightness and contrast, and some features, i. e. skin, were visible in one volume but not in the other.

The results from tables 2 and 3 illustrate that using modern graphics hardware for similarity measure computation can significantly accelerate the process of medical image registration. A hybrid approach, as explained in the previous section, yields results that are as accurate as a CPU-based implementation, in as much as a quarter of the time. The experiments also show that some similarity measures, namely Pattern Intensity and Gradient Difference, deliver very accurate registration even when mipmaps are used for completely summing up the resulting pixels. When registering data sets acquired by using different CT energy levels, similarity measures examining spatial information instead of just the intensity at the individual pixel location deliver very poor results. Still, measures examining pixel intensities alone, especially the Sum of Absolute

Differences, yield a very accurate registration. The centerpiece of the approach, which enabled using the GPUs conveniently for registration of volumetric data has been the projection based method.

## 4 Summary

In this paper, we presented a projection based approach for registration of medical images. Successive registration of two dimensional projection images can yield an accurate volumetric registration. The dimension reduction property of this method, enables convenient usage of GPU for both computing the projections and furthermore estimating the two dimensional similarity images. A series of comparative studies are presented to demonstrate both the accuracy and computational efficiency of the proposed method.

## References

1. Maintz, J., Viergever, M.: A survey of medical image registration. Medical Image Analysis **2** (1998) 1–36
2. Hill, D.L.G., Hawkes, D.J., et. al.: Registration of MR and CT images for skull base surgery using pointlike anatomical features. British Journal of Radiology **64** (1991)
3. Colignon, A., et al.: Automated multi-modality image registration based on information theory, IPMI (1995) 263–274
4. Wells, W., Viola, P., et al.: Multi-modal volume registration by maximization of mutual information. Medical Image Analysis **1** (1996) 32–52
5. Cain, S.C., Hayat, M.M., Armstrong, E.E.: Projection-based image registration in the presence of fixed-pattern noise. IEEE Transactions on Image Processing **10** (2001) 1860–1872
6. Chan, H., Chung, A.C.S.: Efficient 3D-3D vascular registration based on multiple orthogonal 2D projections, Second International Workshop on Biomedical Image Registration (WBIR) (2003) 301–310
7. Lengyel, J., Reichert, M., Donald, B.R., Greenberg, D.: Real-time robot motion planning using rasterizing computer graphics hardware, SIGGRAPH (1990) 327–335
8. Weiskopf, D., Hopf, M., Ertl, T.: Hardwareaccelerated visualization of time-varying 2D and 3D vector fields by texture advection via programmable per-pixel operations, VMV (2001) 439–446
9. Rumpf, M., Strzodka, R.: Level segmentation in graphics hardware. Volume 3., ICIP (2001) 1103–1106
10. Krüger, J., Westermann, R.: Linear algebra operators for GPU implementation of numerical algorithms, SIGGRAPH (2003) 1–9
11. Hajnal, J.V., Hill, D.L., Hawkes, D.J.: Medical Image Registration. CRC Press (2001)
12. Woods, R., Grafton, S., Holmes, C., Cherry, S., Mazziotta, J.: Automated image registration: I. general methods and intrasubject, intramodality validation. Journal of Computer Assisted Tomography **16** (1992) 620–633

13. Weese, J., Buzug, T., Lorenz, C., Fassnacht, C.: An approach to 2D/3D registration of a vertebra in 2D X-ray fluoroscopies with 3D CT images. In: Processding CVRMed/MRCAS. (1997) 119–128
14. Penney, G.P., Weese, J., Little, J.A., Desmedt, P., Hill, D.L.G., Hawkes, D.J.: A comparison of similarity measures for use in 2D-3D medical image registration. IEEE Transactions on Medical Imaging (1998) 586–595
15. LaRose, D.: Iterative X-ray/CT Registration Using Accelerated Volume Rendering. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (2001)

# Combining Homogenization and Registration

Jan Modersitzki and Stefan Wirtz

Institute of Mathematics, University of Lübeck, Germany
{modersitzki, wirtz}@math.uni-luebeck.de

**Abstract.** We present a novel approach for a combined homogenization and registration technique. Medical image data is often disturbed by inhomogeneities from staining, illumination or attenuation. State-of-the-art approaches tackle homogenization and registration separately. Our new method attacks both troublemakers simultaneously. It is modeled as a minimization problem of a functional consisting of a distance measure and regularizers for the displacement field and the grayscale correction term. The simultaneous homogenization and registration enables an automatic correction of gray values and improves the local contrast. The combined approach takes slightly more computing time for an optimization step as compared to the non-combined scheme and so is much faster than sequential methods. We tested the performance both on academic and real life data. It turned out, that the combined approach enhances image quality, especially the visibility of slightly differentiable structures.

## 1 Introduction

A generic task in modern image processing is the integration and/or comparison of data obtained from different images. Particularly in a medical environment, there is a huge demand for comparing pre- and post-intervention images, integrating modalities like anatomy (obtained, e.g., from computer tomography) and functionality (obtained, e.g., from positron emission tomography), or reconstruction of two-dimensional projections to a three-dimensional volume (applies to all tomography techniques and histology).

This task is often disturbed by two sources. The first source is related to a geometrical change of the displayed objects. This change might be introduced, e.g., by motion and/or positioning of objects, distortions and/or changes of imaging devices etc. A well-known technique to compensate these types of distortion is the so-called *image registration*, see, e.g., [5, 7] and references therein. The second source is related to a different appearance of the displayed objects. In histology, e.g., the section typically shows inhomogeneities that are completely related to staining. As a consequence, even after a perfect registration one still observes differences of the displayed objects. While for some applications, e.g., the integration of different modalities, these differences are more or less the goal of the image processing, for other applications, like histology or ultra-sound, these

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 257–263, 2006.

differences are artificial. Typically, *homogenization* is used to compensate these disturbances.

The state-of-the-art approach to the problem is to separate homogenization and registration and to tackle them independently. The homogenization, for example, might be treated based on morphological operations, illumination models, and/or statistical approaches (Markov random fields). For the registration part, there exists a variety of different and well-understood similarity measures in image registration like, for example the sum of squared differences, cross-correlation, or mutual information [2, 11, 9, 7]. All these measures focus on the maintenance of grayvalues and hence none of these measures allow for a correction of grayvalues. There are also some heuristic based schemes for a combined approach. However, these are based one a limited parametric model. A typical example is a straightforward globally linear model, which allows for a mean adaption and a contrast correction [8].

In this paper, we present a novel approach for a combined homogenization and registration technique. Our new approach is based on a sound mathematical formulation that attacks both troublemakers simultaneously. The new method is attractive for grayvalue correction in a variety of applications, including histological serial sectioning (staining), optical flow (illumination), digital radiography (attenuation), and magnetic resonance images (device/user dependency).

The paper is organized as follows. In Section 2 we introduce our notation and give a mathematical foundation of the new combined approach. In Section 3, we describe particular homogenization approaches. Some implementation issues are discussed in Section 4, results for academical as well as real life examples are given in Section 5, and a discussion is given in Section 6.

## 2 The Theoretical Setup

For convenience, we prefer a continuous formulation of the problem. Therefore, our reference $R$ and template $T$ images are mappings from $\Omega \subset \mathbb{R}^d$ to the real numbers $\mathbb{R}$. The mathematical framework covers all image dimensionalities $d$. However, since our main interest is in histology, we focus on $d = 2$.

A standard distance measure in image registration is the $L_2$-norm of the difference image, also called *sum of square differences*, which is basically the energy of the difference image,

$$D(R,T) := \|T - R\|^2 := \int_\Omega \big(T(x) - R(x)\big)^2 \, dx, \tag{1}$$

cf., e.g., [1, 7]. Introducing the deformation $u$ and the deformed template

$$T[u], \quad \text{where} \quad T[u](x) := T(x + u(x)), \tag{2}$$

the basic registration goal is to minimize $D(R, T[u])$ with respect to the geometry $u$. Note that this problem is ill-posed [3, 7] and thus needs to be regularized. Since this regularizer is not central to this paper, for the sake of simplicity, we only focus on the elastic potential $S$, see [7] for details and further regularization. However, our approach has no limitations to this particular regularization.

In order to compensate inhomogeneities, we introduce a correction function $c$ which in addition enables a multiplicative gray value change of the deformed template

$$cT[u], \quad \text{where} \quad (cT[u])(x) := c(x) \cdot T(x + u(x)) \tag{3}$$

and our final distance measure $\|cT[u] - R\|^2$ is to be minimized with respect to the geometry $u$ and the homogeneity correction $c$. For the particular choice $c \equiv 1$, we get the "standard" registration problem back. Setting $c(x) := 1$ for all $x$ where $T(x) = 0$ and $c(x) := R(x)/T(x)$ otherwise, we obtain a trivial minimizer without any geometrical correction. It is obvious that this solution is not helpful. Thus, an additional regularizer $H$ for $c$ has to be introduced. We discuss choices in the next section. The final formulation of the registration problem reads

$$J(u, c) := \|cT[u] - R\|^2 + \alpha S[u] + \beta H[c] \stackrel{!}{=} \min, \tag{4}$$

where $\alpha$ and $\beta$ are regularization parameters and the minimization is with respect to $u$ and $c$ simultaneously. It is very important, that this formulation already combines the minimization of geometry and homogeneity.

## 3 Homogeneity Correction

The central and remaining question is, how to regularize the homogeneity correction $c$? The answer to this question is related to the variation of $c$. If $c$ varies to much (e.g., $c = R/T$), the solution of (4) is meaningless. If on the other hand $c$ does no vary (e.g., $c = 1$) it will not compensate inhomogeneities. We therefore propose a gradient based regularizer,

$$H_p[c] := \int_\Omega \|\nabla c\|^p dx.$$

Obvious choices for $p$ are $p = 2$ (diffusivity) or $p = 1$ (total variation). Simple tests show that the first choice leads to blurred images $cT[u]$. We thus prefer the total variation approach, since it leads to piecewise continuous corrections.

Alternatively, one could also use an explicit regularized version. For example, in histology we might change mean gray value and contrast by a parametric model

$$(cT)(x) = \begin{cases} \gamma_1 T(x) + \gamma_2, & T(x) \neq 0 \\ 0, \text{else} \end{cases}.$$

Here, $c$ is parameterized by $\gamma_1$ and $\gamma_2$ and varies basically over linear maps. Due to this limitations, no additional regularization is needed and we may set $\beta = 0$ in (4).

## 4 Implementation

The new approach is based on a variational formulation of the combined registration problem. From this mathematically sound approach, we derive Euler-Lagrange equations as a system of necessary conditions for a minimizer. The

efficient solution of these partial differential equations is the backbone of our algorithm. However, any adequate numerical scheme can be used.

The Euler-Lagrange equations for (4) are given by

$$J_u = 2c(cT - R)\nabla T - 2\alpha(\mu\Delta u + (\mu + \lambda)\nabla\nabla \cdot u) = 0, \tag{5a}$$

$$J_c = (cT - R)T - \nabla \cdot \tfrac{\nabla c}{\|\nabla c\|} = 0. \tag{5b}$$

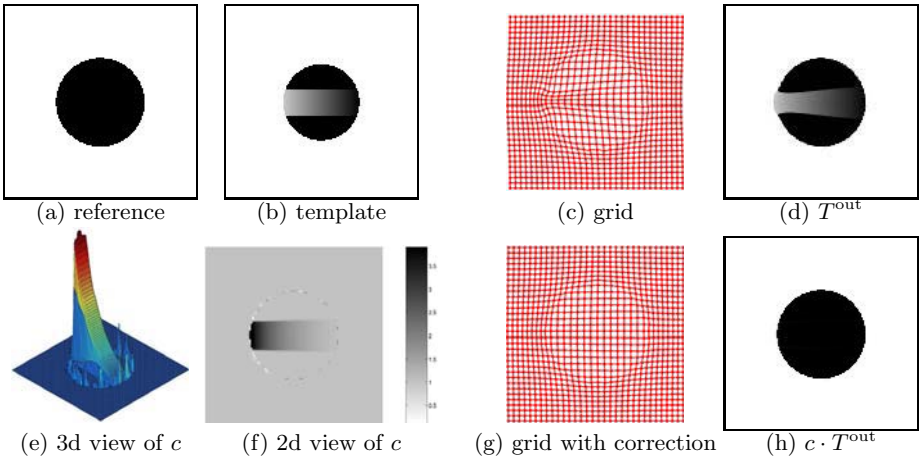
This system of non-linear partial differential equations is discretized using standard finite differences. In principle, a fast multigrid solver has to be used in order to solve these equations efficiently. However, a proper multigrid treatment of total variation - particularly if combined with the elastic operator - is non-trivial and topic of current research and a forthcoming paper. Here, the particular solution scheme is minor and we thus take the solution scheme based on [4] and a straightforward solve for the total variation. Note that our limitation to periodic boundary conditions for the displacement is only due to this particular solution scheme but not part of our method. For a solution of the discrete version we use a non-linear Gauss-Seidel approach. The implementation is coded under MATLAB [6] and executed on an AMD Athlon XP 2700+ with 2 GB RAM. The overall computation time is about one minute for the academical and about three minutes for the real life data. Note that this code runs completely under MATLAB and is far away from being optimized or tuned. We basically aim to show the ability of the new method rather than to focus on a smart implementation.

## 5 Results

We implemented our new scheme and tested the performance on a variety of different examples ranging from academical to real life data. For the academical examples, it is obvious, that the new method benefits from the combined approach. This manifests not only in a much faster execution time, but most importantly in significantly improved registration results. It is easy to construct examples where inhomogeneous regions lead to miss-registration of non-simultaneous schemes. Particularly, we present detailed results for an academical and a real life example.

We tuned our academical example in order to emphasize the power of the simultaneous approach. Our reference image displays a disk, our template image a smaller disk with a grayscale ramp in its interior. Without a simultaneous approach, the inner ramp immediately leads to displacements within these regions; see Figure 1c. With the simultaneous approach, we are not only able to compute a reasonable displacement field (Figure 1g) but also a grayscale correction (Figure 1e and 1f). Only due to this correction we are enabled to display the final $cT[u]$, which in this example happens to be a perfect copy of the reference image. Parameter used: $\alpha = 2e5$, $\beta = 1e3$, $\lambda = 0$, $\mu = 1$, ten iterations (for the simultaneous approach, the minimizer was already obtained after 4 iterations).

Examples for a real life application are also presented. Here we show the results of a registration of two images from consecutive sections within a serial



(a) reference (b) template (c) grid (d) $T^{\text{out}}$

(e) 3d view of $c$ (f) 2d view of $c$ (g) grid with correction (h) $c \cdot T^{\text{out}}$

**Fig. 1.** Academic example: (a) disk as reference, (b) smaller disk with ramp as template, (c) deformed grid (no grayscale correction), (d) deformed template (no grayscale correction), (e) 3d and (f) 2d visualizations of the homogenization correction, (f) deformed grid (with grayscale correction), (g) deformed and corrected template

sectioning of about 800 sections; see [10] for an overview on the overall procedure. As expected, the grayscale correction is not as important as in the academical example. This can be observed from the two deformed grids (Figure 2e and 2h) which more or less look the same. However, a staining trend in this data (visible in the difference image Figure 2f) has been resolved with our combined approach. In contrast to the plain scheme, the combined registration approach results in a visually more pleasing and anatomically superior three-dimensional reconstruction. Parameter used: $\alpha = 5e5$, $\beta = 5e4$, $\lambda = 0$, $\mu = 1$, 40 iterations.

## 6 Discussion

We present a novel registration scheme which enables an unified and simultaneous treatment of the combined registration problem. Our numerical experiments indicate, that the new approach outperforms the uncombined scheme. Our academical examples indicate that the results obtained from the combined scheme are much more reliable. This also applied for the histological serial sectioning, however, the phenomena are less pronounced. Still and most importantly, the automatic homogenization enables an automatic correction of gray values and improves the local contrast. Thus, it leads to a much better visualization of otherwise non-differentiable structures.

Based on an additional solution step for the grayscale correction, the combined scheme takes slightly more computing time for an optimization step as compared to the non-combined scheme. However, this disadvantage is compensated by the fact, that the registration and grayscale correction results are obtained

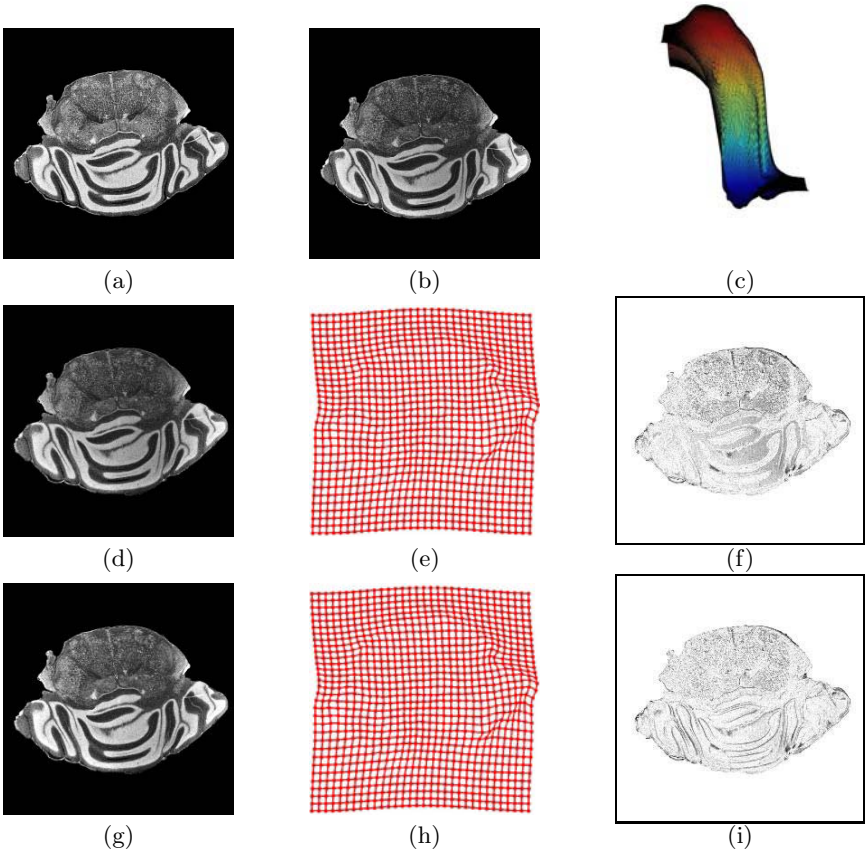



**Fig. 2.** Histological serial section: (a) reference, (b) template: consecutive section, (c) 3d visualization of the homogenization, (d) deformed template (no grayscale correction), (e) deformed grid (no grayscale correction), (f) difference image (no grayscale correction) (g) deformed template (with grayscale correction), (h) deformed grid (with grayscale correction), (i) difference image (with grayscale correction)

simultaneously. A faster and more efficient implementation of the combined approach is topic of further research. In addition, we like to adapt the combined approach to other areas of applications.

## References

1. Lisa Gottesfeld Brown, *A survey of image registration techniques*, ACM Computing Surveys **24** (1992), no. 4, 325–376.
2. A. Collignon, A. Vandermeulen, P. Suetens, and G. Marchal, *3d multi-modality medical image registration based on information theory*, Kluwer Academic Publishers: Computational Imaging and Vision **3** (1995), 263–274.
3. M. Droske and M. Rumpf, *A variational approach to non-rigid morphological registration*, SIAM Appl. Math. **64** (2004), no. 2, 668–687.

4. B. Fischer and J. Modersitzki, *Fast inversion of matrices arising in image processing*, Num. Algo. **22** (1999), 1–11.
5. J Hajnal, D Hawkes, and D Hill, *Medical image registration*, CRC Press, 2001.
6. MathWorks, Natick, Mass., *Matlab user's guide*, 1992.
7. J. Modersitzki, *Numerical methods for image registration*, Oxford University Press, 2004.
8. J. Modersitzki and O. Schmitt, *Image registration of histological sections*, Preprint A-02-06, Institute of Mathematics, University of Lübeck, Germany, 2002.
9. A. Roche, *Recalage d'images médicales par inférence statistique*, Ph.D. thesis, Université de Nice, Sophia-Antipolis, France, 2001.
10. O. Schmitt and J. Modersitzki, *Registrierung einer hochaufgelösten histologischen Schnittserie eines Rattenhirns*, Bildverarbeitung für die Medizin 2001 (H Handels, A Horsch, TM Lehmann, and HPH Meinzer, eds.), Springer, 2001, pp. 174–178.
11. Paul A. Viola, *Alignment by maximization of mutual information*, Ph.D. thesis, Massachusetts Institute of Technology, 1995, pp. 1–155.

# High-Dimensional Normalized Mutual Information for Image Registration Using Random Lines

A. Bardera, M. Feixas, I. Boada, and M. Sbert

Institut d'Informàtica i Aplicacions, Universitat de Girona, Spain
{anton.bardera, miquel.feixas, imma.boada, mateu.sbert}@udg.es

**Abstract.** Mutual information has been successfully used as an effective similarity measure for multimodal image registration. However, a drawback of the standard mutual information-based computation is that the joint histogram is only calculated from the correspondence between individual voxels in the two images. In this paper, the normalized mutual information measure is extended to consider the correspondence between voxel blocks in multimodal rigid registration. The ambiguity and high-dimensionality that appears when dealing with the voxel neighborhood is solved using uniformly distributed random lines and reducing the number of bins of the images. Experimental results show a significant improvement with respect to the standard normalized mutual information.

## 1 Introduction

Multimodal image registration is an important component of medical image analysis used to match two images. It consists in finding the transformation that brings one image into the best possible spatial correspondence with the other one. A common method of solving the registration task is to treat it as a mathematical optimization problem, using a similarity measure to quantify the quality of the alignment of the two images for any given transformation.

Some information-theoretic measures, such as *mutual information* (MI) [1, 2] and *normalized mutual information* (NMI) [3], have become a standard reference due to their accuracy and robustness. However, MI-based methods have a number of well-known drawbacks, such as grid [4] and interpolation effects [1, 5]. Another limitation of these methods is that the computation of the joint histogram is calculated from the correspondence between individual voxels in the two images. In recent years, different approaches have considered a region-based correlation to compute image similarity [6, 7, 8, 9, 10].

In this paper, we propose a new MI-based framework that uses structural information in an image. The NMI measure is extended to consider the correspondence between regions of voxels in multimodal rigid registration. The problems that appear when dealing with the voxel neighbourhood are tackled using uniformly distributed random lines. Experimental results analyze the behaviour of our approach when neighbour intensity values are considered.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 264–271, 2006.

This paper is organized as follows. In Section 2 we survey background and related work, and in Section 3 we present our framework for image registration. Experimental results are given in Section 4.

## 2 Background and Related Work

In this section we review some basic information-theoretic measures [11, 12], the application of mutual information to image registration [1, 2, 13] and recent approaches which take into account spatial information.

### 2.1 Information-Theoretic Measures

Let $\mathcal{X}$ be a finite set, let $X$ be a random variable taking values $x$ in $\mathcal{X}$ with distribution $p(x) = Pr[X = x]$. Likewise, let $Y$ be a random variable taking values $y$ in $\mathcal{Y}$. The *Shannon entropy* $H(X)$ of a random variable $X$ is defined by $H(X) = -\sum_{x\in\mathcal{X}} p(x)\log p(x)$ and measures the average uncertainty of random variable $X$. If the logarithms are taken in base 2, entropy is expressed in bits. The *conditional entropy* is defined by $H(X|Y) = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} p(x,y)\log p(x|y)$, where $p(x,y) = Pr[X = x, Y = y]$ is the joint probability and $p(x|y) = Pr[X = x|Y = y]$ is the conditional probability. The conditional entropy $H(X|Y)$ measures the average uncertainty associated with $X$ if we know the outcome of $Y$. The *joint entropy* $H(X,Y)$ is given by the Shannon entropy of the joint probability distribution. The *mutual information* between $X$ and $Y$ is defined by $I(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$ and measures the shared information between $X$ and $Y$.

We review now the definition of block entropy and entropy rate. The notation used here is inspired by the work of Feldman and Crutchfield [12]. Given a chain $\ldots X_{-2}X_{-1}X_0X_1X_2\ldots$ of random variables $X_i$ taking values in $\mathcal{X}$, a block of $L$ consecutive random variables is denoted by $X^L = X_1\ldots X_L$. The probability that the particular $L$-block $x^L$ occurs is denoted by $p(x^L)$. The Shannon entropy of length-$L$ sequences or *$L$-block entropy* is defined by

$$H(X^L) = -\sum_{x^L\in\mathcal{X}^L} p(x^L)\log p(x^L), \tag{1}$$

where the sum runs over all possible $L$-blocks. The *entropy rate* is defined by

$$h^x = \lim_{L\to\infty}\frac{H(X^L)}{L}, \tag{2}$$

and it can be written as $h^x = \lim_{L\to\infty} h^x(L)$, where $h^x(L) = H(X_L|X_{L-1}\ldots X_1)$ is the entropy of a symbol conditioned on a block of $L-1$ adjacent symbols. The entropy rate of a sequence measures the average amount of information per symbol $x$ and the optimal achievement for any possible compression algorithm [11].

### 2.2 MI-Based Image Registration

The crucial point of image registration is the choice of a metric. The registration between two images $X$ and $Y$ can be represented by an information channel $X \rightarrow Y$, where its marginal and joint probability distributions are obtained by simple normalization of the corresponding intensity histograms of the overlap area of both images [1]. The most successful automatic image registration methods are based on the MI maximization. This method, almost simultaneously introduced by Maes et al. [1] and Viola et al. [2], is based on the conjecture that the correct registration corresponds to the maximum MI between the overlap areas of the two images. Later, Studholme et al. [3] proposed a normalization of MI defined by

$$NMI(X,Y) = \frac{I(X,Y)}{H(X,Y)} = 1 - \frac{H(X|Y) + H(Y|X)}{H(X,Y)}, \tag{3}$$

which is more robust than MI, due to its greater independence of the overlap area. The necessity of normalization is theoretically justified in [14]. It should be noted that $1 - NMI$ is a true distance (see [14, 15]).

Standard MI-based measures ignore the spatial information contained in the images. Recent research has been focused on overcoming this problem. Rueckert et al. [6] propose to use a second-order MI to incorporate spatial information. The neighbourhood has been defined by the six nearest neighbours of each pixel. Pluim et al. [7] include spatial information by combining MI with a term based on the image gradient of the images to be registered. The gradient term seeks to align locations of high gradient magnitude and similar orientations of the gradients at these locations. Sabuncu and Ramadge [8] include spatial information in the MI-based approach by using spatial feature vectors obtained from the images and use a minimum spanning tree algorithm to estimate the conditional entropy in higher dimensions. Russakoff et al. [9] propose an MI extension which takes neighbourhood regions of corresponding pixels into account. They assume that the high-dimensional distribution is approximately normally distributed. Holden et al. [10] use the derivatives of gaussian scale space to provide structural information in the form of a feature vector for each voxel.

## 3 Method

In this section, we define the high-dimensional normalized mutual information and its computation using uniformly distributed random lines.

### 3.1 High-Dimensional Normalized Mutual Information

The $L$-dimensional normalized mutual information of an $L$-dimensional channel $X^L \rightarrow Y^L$ is defined by

$$NMI(X^L,Y^L) = \frac{I(X^L,Y^L)}{H(X^L,Y^L)} = \frac{H(X^L) + H(Y^L) - H(X^L,Y^L)}{H(X^L,Y^L)}, \tag{4}$$

where the joint entropy $H(X^L,Y^L)$ is given by

$$H(X^L, Y^L) = -\sum_{x^L \in \mathcal{X}^L, y^L \in \mathcal{Y}^L} p(x^L, y^L) \log p(x^L, y^L), \tag{5}$$

$I(X^L, Y^L)$ is the $L$-dimensional mutual information, and $p(x^L, y^L)$ is the joint probability of the $L$-dimensional channel. Now, $x^L$ stands for an $L$-block of intensity values in a given 3D region.

To compute the marginal and joint probabilities, how is $x^L$ selected?, i.e., how is the neighbourhood ambiguity solved? With respect to this problem, different templates have been proposed for two dimensions [12], but the generalization from 1-block approach to $L$-block approaches is a difficult problem.

Another aspect to be considered is the size of an $L$-block. As $L$ increases, $\frac{H(X^L)}{L}$ decreases, converging to the entropy rate $h^x$ of an image, which expresses its maximum compressibility or, equivalently, its irreducible randomness. In fact, the entropy of an image appears more random than it actually is, and the difference is given by $h^x(L) - h^x$ [12]. Therefore, the higher the $L$ value, the better the approximation of $I(X^L, Y^L)$ to the true shared information. On the other hand, note that the number of elements of the joint histogram is given by $N^{2L}$, where $N$ is the number of bins of an image.

From these limitations, how to compute $NMI(X^L, Y^L)$ for image registration? Rueckert et al. [6] propose a second-order MI using the six nearest neighbours of each voxel and reducing the number of bins to 16. One drawback of this approach is that not all directions are taken into account. In the next section, we tackle this problem using uniformly distributed random lines.

### 3.2 Implementation Using Uniformly Distributed Random Lines

Uniformly distributed random lines, also called *global lines* [16], can be utilized to compute the joint intensity histogram in 3D-image registration [17], which is the most demanding step in the similarity measure computation. The overlapping volume between two 3D-images is stochastically sampled using a uniform distribution of lines in the sense of integral geometry, i.e., invariant to translations and rotations [18]. Points chosen on each line provide us with the intensity values to calculate the probability distributions.

A global line can be generated from the walls of a convex bounding box containing the data volume [19]. This can be done taking a random point on the surface of the convex bounding box and a cosinus distributed random direction. This strategy is easily applicable to image registration since the reference image can be taken as the bounding box (Figure 1). The intensity values are captured from the lines at evenly spaced positions, taking an initial random offset different for each line. The random offset ranges from 0 to the step size. Although we skip with regular steps, the use of a random offset ensures the stochasticity of the process. This fact diminishes the interpolation artifacts (see [17]).

In order to implement the $NMI(X^L, Y^L)$ measure, we propose to take $L$-blocks on the global lines (Figure 1). Thus, the 3D-neighbourhood problem is reduced to a 1D problem. Now, the ambiguity about how to sample the neighbourhood disappears as the order in which we take the neighbours is well defined

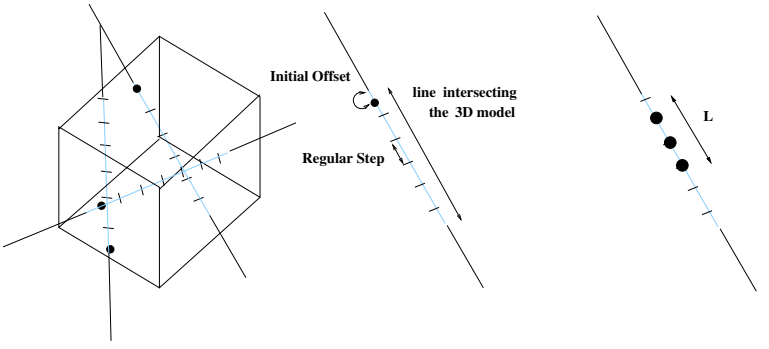



**Fig. 1.** Global lines are cast from the walls of the bounding box. Neighbour intensity values are taken in L-blocks.

on a global line. In spite of this, the dimensionality problem persists due to the fact that the computation of the joint histogram of the channel $X^L \rightarrow Y^L$ has an excessive cost if $L \geq 2$. Another problem is the sparsity of the resulting histograms since the number of bins of the joint histogram is $N^{2L}$. A possible solution to these problems is to reduce the number of bins of the image, preserving a trade-off between level of binning and the accuracy of the entropy rate.

## 4 Results

In this section, two sets of experiments are carried out. First, we analyze the behaviour of $NMI(X^L, Y^L)$ when two neighbour intensity values are considered on random lines and, second, we show preliminary results taking 3-blocks. The clinical images and the standard transformations used for the tests are provided as part of the Retrospective Image Registration Evaluation project [20].

The first experiments use the ITK MI_Validation application [21] with a multiresolution optimization approach using the quaternion rigid transform gradient descent. In our implementation, three multiresolution steps, with 1500 iterations each one, are used. To reduce the computational cost, all the measures are computed using global lines, with a total number of 100000 sampling points and an step size of 2 $mm$. The $NMI$ and $NMI(X^2, Y^2)$ have been computed, respectively, over a range of 256 and 16 bins per dimension. The results of CT-MR_T1 image registrations for 16 patients are shown in Table 1. Each entry in the table is the mean of the distances in $mm$. between the evaluated registration method and the gold standard measured at different volumes of interest (VOI) of the patient. Note that in 14 out of 16 tested patients the proposed approach behaves better than the $NMI$ method, and, in all the cases, the error of our method is lower than 2.5 $mm$., which reveals its good performance. This good behaviour is also shown in Table 2, where the mean, the median and the maximum of the distances between VOIs have been computed from all patient results.

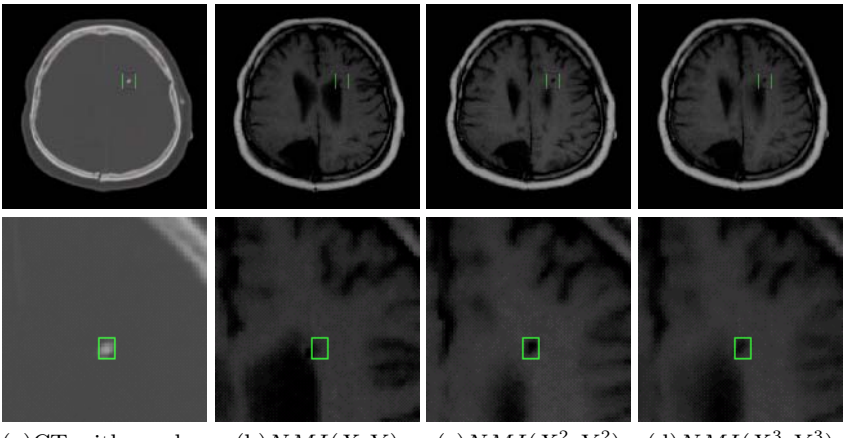
In Figure 2, we also present a preliminary evaluation of the proposed algorithm with blocks of three neighbours. Our method is now used to register the

**Table 1.** The mean of the error in *mm*. for NMI and $NMI(X^2, Y^2)$, measured at different VOIs for each patient

| CT-MR_T1 | $NMI$ | $NMI(X^2, Y^2)$ | CT-MR_T1 | $NMI$ | $NMI(X^2, Y^2)$ |
|---|---|---|---|---|---|
| pat_001 | 2.05 | 0.97 | pat_102 | 5.57 | 1.54 |
| pat_002 | 1.98 | 1.28 | pat_103 | 2.28 | 2.44 |
| pat_003 | 1.78 | 1.10 | pat_104 | 2.91 | 1.22 |
| pat_004 | 13.16 | 2.21 | pat_105 | 11.49 | 2.45 |
| pat_005 | 1.63 | 0.84 | pat_106 | 5.81 | 1.74 |
| pat_006 | 12.62 | 1.44 | pat_107 | 4.18 | 0.94 |
| pat_007 | 0.78 | 1.13 | pat_108 | 9.21 | 1.69 |
| pat_101 | 6.57 | 2.10 | pat_109 | 1.27 | 1.15 |

**Table 2.** Mean, median and maximum of the distances in *mm*. between VOIs from all patient results

| Method | Mean | Median | Maximum |
|---|---|---|---|
| $NMI$ | 5.02 | 3.00 | 19.74 |
| $NMI(X^2, Y^2)$ | 1.50 | 1.31 | 2.92 |



(a)CT with marker (b)$NMI(X, Y)$ (c)$NMI(X^2, Y^2)$ (d)$NMI(X^3, Y^3)$

**Fig. 2.** CT-MR registration results for different $L$ values

CT-MR_T1 pair of the patient 3 of the Vanderbilt database, considering different L values. To overcome the high-dimensionality of the joint histogram, the number of bins has been reduced to 16 bins and 8 bins for $L = 2$ and $L = 3$, respectively. In Figure 2, we show the original CT with a rectangular marker and the transformed MR obtained with L=1, L=2 and L=3 and with the marker in the same geometrical position as in the CT image. In the second row of Figure 2 we show a zoom of the first row images. In this test, the correctness of the registration for each case is determined by the degree of coincidence between the white dot in the center of the marker of the CT image and the black dot in the

MR one. Observe that, for $NMI(X^2, Y^2)$ and $NMI(X^3, Y^3)$, the registration achieves more accurate results than the standard method NMI. This behaviour encourages us to explore the proposed measure with higher L values.

## 5 Conclusions and Future Work

In recent years, mutual information and its normalized versions have emerged as effective similarity measures for image registration. However, one significant drawback is that they ignore structural information in images. In this paper, we have presented the high-dimensional normalized mutual information which takes into account the dependence between image regions. The high dimensionality of the problem has been tackled using uniformly distributed random lines and reducing the number of image bins. Random lines permit us to solve the ambiguity in taking the neighbour intensity values of a voxel and also contributes to the reduction of interpolation artifacts. Experimental results show that an accurate and robust registration is achieved using only two neighbour intensity values in the normalized mutual information computation. In our future work we will analyze the behaviour of our proposal using three or four neighbours combined with a previous non-uniform quantization of the images.

## Acknowledgments

The images used in our experiments were provided as part of the project, "Evaluation of Retrospective Image Registration", National Institutes of Health, Project Number 1 R01 NS33926-01, Principal Investigator Prof. J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN. This project has been funded in part with grant numbers TIN2004-08065-C02-02, TIN2004-07451-C03-01 and 2001-SGR-00296.

## References

1. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging **16**(2) (1997) 187–198
2. Viola, P.A.: Alignment by Maximization of Mutual Information. PhD thesis, Massachusetts Institute of Technology, Massachusetts (MA), USA (1995)
3. Studholme, C.: Measures of 3D Medical Image Alignment. PhD thesis, University of London, London, UK (1997)
4. Unser, M., Thévenaz, P.: Stochastic sampling for computing the mutual information of two images. In: Proceedings of the 5th International Workshop on Sampling Theory and Applications (SampTA'03). (2003) 102–109 Held in Strobl, Austria.
5. Tsao, J.: Interpolation artifacts in multimodal image registration based on maximization of mutual information. IEEE Transactions on Medical Imaging **22** (2003) 854–864

6. Rueckert, D., Clarkson, M.J., Hill, D.L.G., Hawkes, D.J.: Non-rigid registration using higher-order mutual information. In: Proc. SPIE Medical Imaging 2000: Image Processing. (2000) 438–447 San Diego.
7. Pluim, J.P., Maintz, J., Viergever, M.A.: Image registration by maximization of combined mutual information and gradient information. IEEE Transactions on Medical Imaging **19**(8) (2000) 809–814
8. Sabuncu, M.R., Ramadge, P.J.: Spatial information in entropy-based image registration. In: Proceedings of 2nd Workshop in Biomedical Image Registration (WBIR'03), LNCS 2717. (2003) 132–141 Held in Philadelphia, USA.
9. Russakoff, D.B., Tomasi, C., Rohlfing, T., C. R. Maurer, J.: Image similarity using mutual information of regions. (In: Proceedings of the 8th European Conference on Computer Vision (ECCV), LNCS 3023) 596–607
10. Holden, M., Griffin, L.D., Hill, D.L.G.: Multi-dimensional mutual information image similarity metrics based on derivatives of linear scale space. In Lovell, B.C., Maeder, A.J., eds.: Proceedings of the APRS Workshop on Digital Image Computing. (2005) 55–60
11. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley Series in Telecommunications (1991)
12. Feldman, D.P., Crutchfield, J.P.: Structural information in two-dimensional patterns: Entropy convergence and excess entropy. (2002)
13. Pluim, J.P., Maintz, J., Viergever, M.: Mutual-information-based registration of medical images: a survey. IEEE Transactions on Medical Imaging **22** (2003) 986–1004
14. Li, M., Chen, X., Li, X., Ma, B., Vitányi, P.: The similarity metric. IEEE Transactions on Information Theory **50**(12) (2004) 3250–3264
15. Kraskov, A., Stögbauer, H., Andrzejak, R.G., Grassberger, P.: Hierarchical clustering based on mutual information. (2003)
16. Sbert, M.: An integral geometry based method for fast form-factor computation. Computer Graphics Forum (Proceedings of Eurographics'93) **12**(3) (1993) 409–420 Held in Barcelona, Spain.
17. Bardera, A., Feixas, M., Boada, I., Sbert, M.: Medical image registration based on random line sampling. In: IEEE International Conference on Image Processing (ICIP'05), Proceedings, Genova, Italy (2005)
18. Santaló, L.A.: Integral Geometry and Geometric Probability. Addison-Wesley, Reading (MA), USA (1976)
19. Castro, F., Martínez, R., Sbert, M.: Quasi-Monte Carlo and extended first shot improvements to the multi-path method. In: Proceedings of Spring Conference on Computer Graphics'98. (1998) 91–102 Held in Budmerice, Slovak Republic.
20. National Institutes of Health: Retrospective Image Registration Evaluation. Vanderbilt University, Nashville (TN), USA. (2003) Project Number 8R01EB002124-03, Principal Investigator J. Michael Fitzpatrick.
21. ITK Insight Toolkit, http://www.itk.org.

# 2D/3D Registration of Neonatal Brain Images

Ewout Vansteenkiste, Jef Vandemeulebroucke, and Wilfried Philips

Ghent University, Department of Telecommunication and Information Processing, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium

**Abstract.** Recently, it is shown the morphology of the white matter, obtained by ultrasound (US) imaging, has prognostic implications for certain brain pathologies. Periventricular Leukomalacia (PVL) is such a pathology diagnosed on infants with a very low birth weight ($< 1500$g). Next to the US images Magnetic Resonance Image (MRI) volumes are commonly used for the inspection of this pathology. Since on both modalities, up to now, we still lack a golden standard for the quantification of the syndrome cross-validation through a multi-modal registration is highly beneficial to the clinical diagnosis. In this article we present a 2D US - 3D MRI registration scheme combining an interactive initialization step, B-spline image interpolation, a mutual information based metric and an evolutionary algorithm optimization scheme.

## 1 Introduction

1.2 % of the newborns are Very Low Birth Weight (VLBW) infants ($\leq 1500$g). The recent increase in survival rate of these VLBW infants has brought along an increasing incidence of neurological sequelae [1]. Here we will focus on one particular brain pathology, called Periventricular Leukomalacia (PVL), which in its milder variant is characterized by deep white matter lesions, also called flaring, adjacent to the lateral ventricles, see Fig. 1. This is the primary indication of further abnormal brain development. Among the survivors of the neonatal period, 50 % is normal at the age of 6, 10-20 % suffer from severe handicaps.

In previous work [1, 2] PVL has already been studied on MRI volumes as well on newborns as on the outcome on later stages. Other work [3, 4] already showed the possibility of detecting PVL based on US images. The non-invasiveness and portability of the US machine make this modality highly suitable given the nature of the patients. On the other hand, due to the speckle noise present, the visual inspection of the images is often subjective.

Although MRI volumes are less noisy, they are harder to capture when dealing with non-sedated preterms. Next to that, we still lack a golden standard for the quantification of the flaring both in US as well as in MRI. Therefore, physicians nowadays believe valuable information can be obtained from the simultaneous inspection of the US and the corresponding MR image. In order to obtain this objective image comparison, the images should contain exactly the same anatomical features and thus have to be registered correctly.

In literature, many 3D/3D registration algorithms have been presented for different modalities, also including US. The 2D/3D case is often less trivial given

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 272–279, 2006.

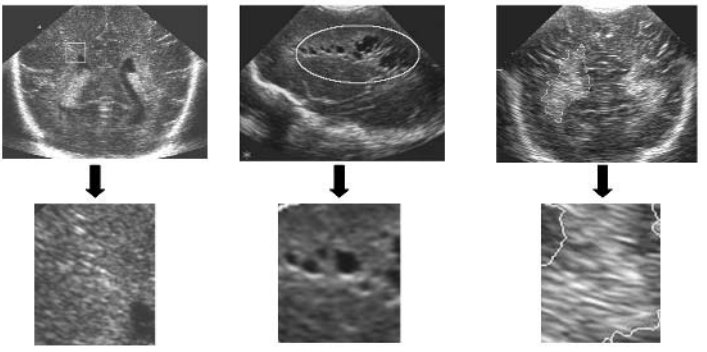


**Fig. 1.** On the left a normal brain and part of the periventricular zone below it. In the middle, a cystyc PVL brain and below it a clear cut-out of the cysts in the tissue that come with the severe variant of the pathology. On the right, the gliotic PVL with the flaring delineated and below a cut-out of the periventricular flaring.

there is less information available for the same number of degrees of freedom. Next to that, 2D/3D schemes are commonly found for CT, PET, MRI modalities, but the typical speckle noise in the US images makes it hard to create this kind of registration scheme, although there is a clear demand for it. Here we will try to fill the gap by presenting an interactive registration scheme, to our knowledge the first on neonatal US brain imaging.

In the next sections we will first explain the data used Section 2, then present our registration scheme, Section 3, show some experimental results in Section 4, before presenting our conclusions in Section 5.

## 2 Data Acquisition and Preprocessing

A total of 20 coronal 2D-US brain images obtained from 20 preterms up to 32 weeks of gestation were analyzed. All images were captured in the first 3 weeks after birth at the Sophia Children's hospital Rotterdam, The Netherlands, by one medical expert using an Acuson Sequoia 512 ultrasound machine with a hand-free 8.5 MHz probe. The US image pixel size is $0.16 \times 0.16$ mm. Consequently, 20 MRI volumes of the same patients were acquired using a General Electric 1.5T Sigma Infinity scanner. A T1 weighted volume containing about 85 images at a voxel size of $0.85 \times 0.85 \times 1$ mm and T2 weighted volume containing about 20 images at a voxel size of $0.70 \times 0.70 \times 4$ mm were used in the registration process. The acquisition time for the MRI images lay between 4 and 5 minutes.

Given the conical shape of the US image there is a lot of information on the image at display that is redundant to the registration procedure. Therefore, instead of working on the raw US images we first cut out all superficial information, being all black background not part of the brain as well as all parts outside of the skull, where the US beams do not penetrate.

Consequently, we incorporate the scan angle of the ultrasound image. Although in this experiment the US images were acquired without any probe tracker, the medical expert scans under a 45 degree angle to the coronal plane. Thus, in rotating the MRI volume over the same angle we obtain a better starting position.

## 3 Registration Scheme

### 3.1 Mechanism

Fig. 2 illustrates our registration scheme, apart from the previously explained initialization steps. We will briefly discuss this scheme before going into detail on the most important parts. First, we have to decide on which image will be fixed in a space coordinate system, and which one will be transformed resulting each time in a new overlap. We choose the 2D-US image as fixed, the MRI

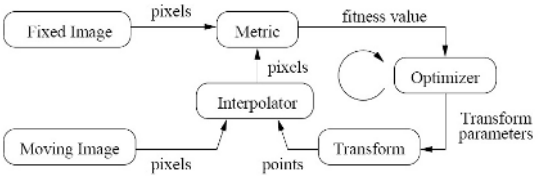



**Fig. 2.** Schematic overview of the registration

volume as moving. For each possible overlap, an interpolator calculates the values of the moving image at the grid positions of the fixed image. On this overlap, we calculate a metric value, which is an adequate indicator for the similarity between the current overlapping (MR) slice and the fixed (US) image. Based on this metric value, an optimizer then proposes new transform parameters. Before we can evaluate this new candidate solution, the interpolator once again finds the intensities of the moving image at the grid positions of the fixed image. This cycle is repeated until the optimal transform is found or a stopping criterium is reached.

More technically, suppose $u$ is our reference image and $v$ our test volume, and $u(x)$ and $v(x)$ are the corresponding pixels or voxels. Denote by $T$ a transformation of the coordinate system of the reference image to that of the test volume and by $F_c$ a function that reaches a maximum when the registration is correct. Then the registration goal is to find the optimal estimate transformation $T_c$, so that

$$T_c = arg \max_T F_c(u(x), v(T(x))). \tag{1}$$

### 3.2 Transform

Since we include 3D volumes, we consider 6 degrees of freedom, allowing translations and rotations around the three principal axes. The head is considered a rigid object so scaling is not applied. The transform is centered, meaning the center of rotation can be chosen arbitrary other than in an origin based implementation. A natural choice is the center of the head, which leads to an

easier interpretation of and a reduced dependency on the transform parameters. If for example all parameters are set to their optimal values, varying one rotation angle will result in the volume rotating, while still remaining in an overlapping position with the US image. A similar movement with a not-centered transform would require changes in all 6 transform parameters. In Fig. 3, a 2-dimensional example is given, illustrating both types of transforms.

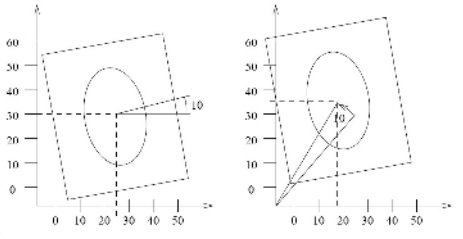


**Fig. 3.** On the left a rotation around a center, on the right a rotation around the origin

### 3.3 Interpolator

The scan time for the MRI volumes has to be limited and we are restricted to work by the protocols put forward by the medical doctors. This implies a non-neglectable difference between the resolution of both modalities. For example, the US images used have a resolution of $0.16 \times 0.16$ mm while the T2 weighted MRI volume has a resolution of $0.7 \times 0.7 \times 4$ mm. Hence we interpolate the MRI volume, to assure acceptable registration results. As interpolating is computationally expensive, it is essential to use a time-efficient algorithm. For this reason a simple linear interpolation is performed in each iteration of the scheme.

For some volumes however, as the T2 weighted volume, it is necessary to perform a more complex interpolation before the actual registration precedure starts, since just applying a linear interpolation would result in images of poor quality. That is why the MRI volume is first upsampled using B-splines of degree 5. In order not to increase the memory usage exponentially, we created isotropic voxels of size 0.5mm in that way. Fig. 4, shows an example of an interpolation using B-splines.

### 3.4 Metric

This component is undoubtedly the most critical. The metric value is the result of a mathematical measure, based on the values of the group of pixels relevant for the current solution and reflects the quality of this solution. Since we are working with different modalities, we can not just compare the intensities of corresponding pixels the two images.

Here we used an implementation of the Mutual Information algorithm provided by Mattes et al. [6] which is computationally more interesting than the one proposed by Viola et al. [5]. For more details we refer to the reference paper, but mainly only one sample set is used throughout the entire run of the program

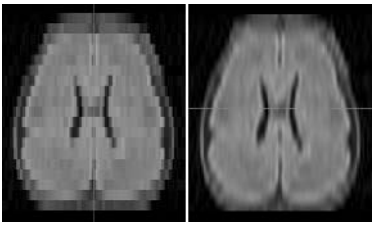


**Fig. 4.** On the left the original MR image (z-direction), on the right the B-spline interpolated MR image

and B-splines, having a finite support, instead of Gaussian kernels are used in order to estimate the probability density functions (PDF). For the calculation of the entropy, no second sample set is needed but the estimated PDF's, uniformly distributed over the intensity interval, are evaluated. The advantage of this metric is that it uses only 2 parameters: the number of samples used to estimate the PDF, and the number of histogram bins used to calculate the entropy.

### 3.5 Optimizer

The optimizer finds the optimal transformation parameters. It searches the 6-dimensional space to efficiently zoom in on the optimal solution, while not getting caught in local extrema of the metric. A $(1 + 1)$-evolutionary strategy was implemented, which has proven to be more robust and effective in this kind of environment.

The strategy starts from a set op possible initial solution and combines them in a specific way as to obtain new generations of solutions. Suppose we obtain a solution $x_t$, at iteration $t$. A new solution $x_{t+1}$ is then generated, satisfying:

$$x_{t+1} = x_t + a_t r_t$$

where $r_t$ is a normally distributed random variable:

$$r_t \sim N(0, 1)$$

and $a_t$ satisfies:

$$a_t = \begin{cases} a_{t-1} c_{grow} & \text{if } F(x_t) > F(x_{t-1}) \\ a_{t-1} c_{shrink} & \text{else.} \end{cases}$$

Note that here we assume a higher value $F(x)$ corresponds to a better solution. The parameters $a_0$, $x_t$, $c_{grow}$, $c_{shrink}$ are set priorly. Here, $x_0$ is an initialization of $T$, the value of $c_{grow}$ en $c_{shrink}$ are set according to [7], and are close to 1. The algorithm stops when a prefined maximum number of generations is reached or if $a_t$ is smaller then a prefined threshold $\varepsilon$.

Intuitively, we impose normal variations on a solution at a certain stage scaled by a factor $a_t$. If we find a better solution than the current one we increase this factor, thus supposing even better solutions may lay even further away, improving the exploration of the search space.

If no better solution is found we decrease this factor, thus supposing a better solution can be found through the exploitation of the local neighborhood. The small number of floating point operations makes the algorithm very fast and allows us to generate many generations without enlarging the computation time significantly. Another advantage of these optimizations is their relative insensitivity to local optima in the metric, allowing the technique to perform well in noisy environments, such as the US images.

The major drawback is that the solution is not always unique and so multiple runs have to be taken into account to base the final result on. As mentioned above the low computational complexity makes this possible.

## 4 Experimental Results

Fig. 5 shows the visual results of the registration of 2 different patients. The mosaic images are shown as well as the corresponding US and MR image. The left images are based on the T1 weighted volume, the right on the T2 weighted volume. All results were obtained using the same protocol. First, the appropriate US region of interest was selected, then the MRI volumes were interpolated using the B-splines, rotated of a 45 degrees angle and an initial starting slide was chosen. As such transform was initialized as

$$T_i = (-0.8,\ 0,\ 0,\ 0,\ \frac{t_z}{\sqrt{2}},\ \frac{t_z}{\sqrt{2}})$$

with $t_z$ the z-coordinate of the initial slide, first 3 coordinates presenting the rotation (in radials), last three the translation.

Instead of using an initial random solution for the evolutionary algorithm, a Regular Step Gradient Descent method combined with the Mattes mutual information was run on $T_i$, leading to the initial starting point for the scheme presented earlier. This was found more stable than any arbitrary selection. The

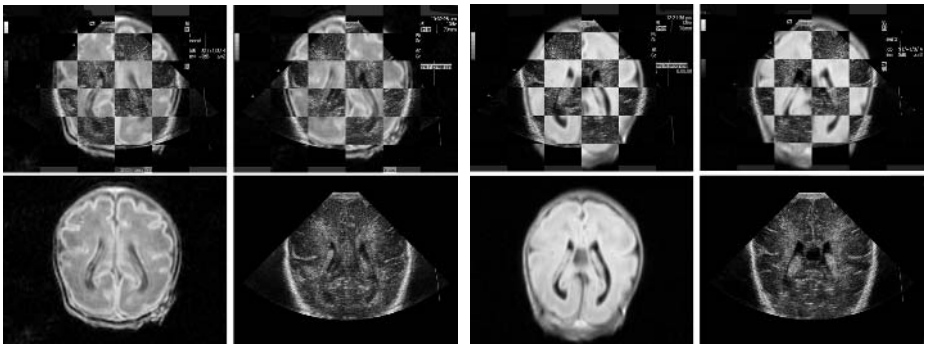


**Fig. 5.** Registration results for 2 different patients, on the left a T1 weighted image, on the right a T2 weighted image

registration procedure was then run 15 times and the optimal registration result was selected based on the results of the metric, next to the visual evaluation of the results. The optimal translation parameters for the two patients can be found in the Table below.

| $Translation$ | $r_x$ | $r_y$ | $r_z$ | $t_x$ | $t_y$ | $t_z$ |
|---|---|---|---|---|---|---|
| $P1_{initial}$ | -0.8 | 0 | 0 | 0 | 11.3 | 11.3 |
| $P1_{final}$ | -0.816 | -0.00772 | -0.00493 | -0.340 | 10.6 | 11.8 |
| $P2_{initial}$ | -0.8 | 0 | 0 | 0 | -16 | -16 |
| $P2_{final}$ | -0.639 | -0.0347 | 0.0505 | 2.71 | -12.4 | -16.8 |

Optimal results differed 0.1 mm in translation and 0.01 radials in rotation compared to the manual registration of a medical expert, which is very good given the very small size of the head. We should mention however that solely based on the results of the metric, the rotational variance might be bigger. Therefore, a visual inspection of the best outcomes is indeed necessary. Note also that the final results often are rather close to the initializations so a good initialization is very important. Given the spline interpolation and multiple runs of the genetic optimizer the whole process also takes quite some computation time, up to 3 hours momently. This we are currently trying to optimize.

When looking at the visual results, also see Fig. 6, we see the outer structures of the brain are usually registered well, there were for the inner parts it is less obvious. This can be explained by the fact that some features such as the Plexus Chorideus, an anatomical region inside of the lateral ventricles, are displayed differently due the intrinsic properties of the different modalities. This makes the US registration very hard and calls for a visual interpretation of the results as mentioned above.

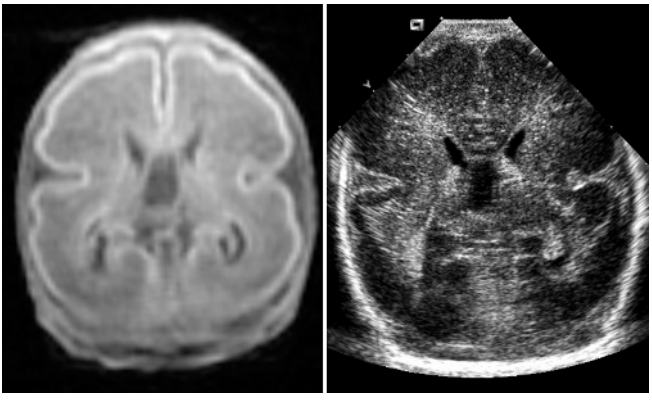


**Fig. 6.** Registered MR and US image. The Plexus Chorideus is the bright white spot at the bottom of each central ventricle. Notice is has, amongst other features, a different appearence on the MR and US image.

Finally no real difference was found in the registration results for the T1 and T2 weighted images, altough the T1 images tend to show more detailed information, making them more suitable for visual diagnosis.

## 5 Conclusions

We constructed a 2D US - 3D MRI registration method with limited user-interaction. The program automatically places both images in the space coordinate system, after upsampling the MRI volume. The user then roughly initializes the transform, after which an evolutionary strategy optimizes a mutual information based metric. Visual results, as shown in Fig. 5, are satisfying and have been validated by the clinical experts.

As a continuation of our work we would like to test how a feature-based metric could even further improve the registration process as well as possible speckle suppression as a preprocessing to the registration algorithm. Based on this registration, first steps are also taken in order to cross-validate the specific segmentation and classification results found on both modalities separately.

## References

1. D. Davis. Review of cerebral palsy. *Neonatal Network*, vol. 16, pp. 7–12, 1997.
2. N. Paneth, R. Nudelli, E. Kazam and W. Monte. Brain Damage in the preterm infant. *Mac Keith Press*, London, 1994.
3. S. Fujimoto, N. Yamaguchi, H. Togari, Y. Wada and K. Yokochi. Cerebral palsy of cystic periventricular leukomalacia in low-birth-weight infants. *Acta Paediatr*, vol. 83, pp. 397–401, 1994.
4. F. Pidcock, L. Graziani, C. Stanley, D. Mitchell and D. Merton. Neurosonographic features of periventricular echodensities associated with cerebral palsy in the preterm infant. *Journal of Pediatrics*, vol. 116, pp. 417–422, 1990.
5. P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *IJCV*, vol.24, no. 2, 137154, 1997.
6. D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. Non-rigid multimodality image registration. *Medical Imaging 2001: Image Processing*, pp. 1609–1620, 2001.
7. M. Styner and G. Gerig. Parametric Estimation of 2D/3D Bias Correction with 1+1ES-Optimization. *Techmical Report*,Image Science Lab ETH Zürich, 1997.

# Robust Optimization Using Disturbance for Image Registration

Rui Gan and Albert C.S. Chung

Lo Kwee-Seong Medical Image Analysis Laboratory,
Department of Computer Science,
The Hong Kong University of Science and Technology, Hong Kong
{raygan, achung}@cs.ust.hk

**Abstract.** This paper exploits the different properties between the local neighborhood of global optimum and those of local optima in image registration optimization. Namely, a global optimum has a larger *capture neighborhood*, in which from any location a monotonic path exists to reach this optimum, than any other local optima. With these properties, we propose a simple and computationally efficient technique using transformation disturbance to assist an optimization algorithm to avoid local optima, and hence to achieve a robust optimization. We demonstrate our method on 3D rigid registrations by using mutual information as similarity measure, and we adopt quaternions to represent rotations for the purpose of the unique and order-independent expression. Randomized registration experiments on four clinical CT and MR-T1 datasets show that the proposed method consistently gives much higher success rates than the conventional multi-resolution mutual information based method. The accuracy of our method is also high.

## 1 Introduction

An important component in the medical imaging field is multi-modal image registration, which can integrate complementary image information acquired from different modalities. The task of image registration is to reliably identify a geometric transformation to accurately align two images. General promising results have shown that mutual information (MI) as a voxel intensity-based similarity measure is well-suited for multi-modal image registration [8, 14]. However, it has been suggested that the conventional mutual information based registration method can result in misalignment for some cases [9, 11]. One possible causation is that the mutual information based method can get trapped into local maxima during transformation optimization process. Most of the existed optimization methods, e.g. Powell's method, simplex methods, gradient descent and so on [12], are only suitable for local optimization. Several attempts have been made to improve the optimization performance, such as multi-resolution approaches [2], extension and combination of different optimization methods [10, 5]. However there is still no guarantee to find the global solution in general. On the other hand, global stochastic optimization methods, e.g. Simulated Annealing [4], need

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 280–288, 2006.

to take a large number of iterations to converge, which limit their application in image registration.

In this paper, in an attempt to obtain a robust optimization result, a simple and computationally efficient technique is proposed to assist an optimization method to avoid local optima. The technique is based on the different properties between the local neighborhood of global optimum and those of local optima. Specifically, the global optimum has a larger *capture neighborhood*, in which from any location a monotonic path exists to reach this optimum, than any other local optima. When an optimization procedure gets trapped into a local optimum, a proper disturbance on the obtained transformation can pull it out of the undesirable location and consequently provide a further chance for a following up optimization to achieve the global optimum. The proposed method is demonstrated on 3D rigid registrations. We adopt quaternions to represent rotations for the purpose of the unique and order-independent expression. Based on the randomized registration experiments on four clinical 3D CT and MR-T1 image volumes, it is demonstrated that the new method consistently gives much higher successful registration rates than the conventional multi-resolution mutual information based method. The results also implies that our method can obtain acceptably high registration accuracy.

## 2 Methods

### 2.1 Mutual Information as Similarity Measure

Mutual information (MI) is an useful concept from information theory [3] and measures the amount of information shared between two random variables. Specifically, mutual information quantifies the Kullback-Leibler distance [7] between the joint distribution of two random variables, A and B, and the product of their marginal distributions, that is

$$MI(A,B) = \sum_{a,b} p(a,b) \log \frac{p(a,b)}{p(a) \cdot p(b)}, \tag{1}$$

where $p(a,b)$ is the joint distribution of A and B, and $p(a)$ and $p(b)$ are the individual marginal distributions respectively.

Mutual information was proposed independently as similarity measure for 3D rigid registration of medical images by Wells et al. [14] and Maes et al. [8]. To utilize mutual information, the intensity values of the corresponding voxel pair in the two images to be registered are considered as random variables, and the joint and marginal distributions can be the normalization of the joint and marginal histograms of the sampling set. The mutual information of two images measures the amount of information conveyed by one image that is shared by the other image, and it is assumed to be maximum when the images are aligned. Therefore, the mutual information based registration method identifies a geometrical transformation $\hat{\mathbf{T}}$ as follows,

$$\hat{\mathbf{T}} = \arg\max_{\mathbf{T}} MI(\mathbf{T}) = \arg\max_{\mathbf{T}} MI(\mathbf{T}(I^f), I^r),$$

where $I^r$ and $I^f$ are the intensity domains for the *reference* and *floating* images respectively.

### 2.2 Robust Optimization Using Transformation Disturbance

General promising results have shown that the mutual information based registration method works well on multi-modal images [8, 14]. However, it has been suggested that it can result in misalignment for some cases [9, 11]. One possible causation is that mutual information based method can get trapped into local maxima during transformation optimization process. This motivates us to propose a technique to assist an optimization method to avoid those local maxima.

From any initial transformation $\mathbf{T}_0$, it is obvious that there exists a *monotonic path*, $\overrightarrow{\mathbf{T}_0\mathbf{T}_1\cdots\mathbf{T}_l}$, to a local (or global) maximum $\mathbf{T}_l$, such that $MI(\mathbf{T}_0) \leq MI(\mathbf{T}_1) \leq \cdots \leq MI(\mathbf{T}_l)$. Then, the *capture neighborhood* of an arbitrary local (or global) maximum $\mathbf{T}_{l_i}$, $\mathcal{N}_{\mathbf{T}_{l_i}}$, can be defined as a maximum local neighborhood in which any transformation has such a monotonic path to reach $\mathbf{T}_{l_i}$. Alternatively, $\mathcal{N}_{\mathbf{T}_{l_i}}$ can be expressed as follows,

$$\mathcal{N}_{\mathbf{T}_{l_i}} = \{\mathbf{T} \mid \exists\ \overrightarrow{\mathbf{T}\cdots\mathbf{T}_{l_i}},\ \text{such that}\ MI(\mathbf{T}) \leq \cdots \leq MI(\mathbf{T}_{l_i})\},$$

and its radius is given by

$$R_{\mathbf{T}_{l_i}} = \max_{\mathbf{T}\in\mathcal{N}_{\mathbf{T}_{l_i}}} d(\mathbf{T}, \mathbf{T}_{l_i}),$$

with $d(\cdot)$ being the distance of two locations. With this definition, the whole transformation space can be treated as the union of all such capture neighborhoods, i.e. $\bigcup_i \mathcal{N}_{\mathbf{T}_{l_i}}$.

The foundation of the proposed method is an observation that the radius of the capture neighborhood of an arbitrary local maximum $\mathbf{T}_{l_i}$ is much smaller than that of the global maximum $\mathbf{T}_o$, i.e. $R_{\mathbf{T}_{l_i}} \ll R_{\mathbf{T}_o}, \forall\ \mathbf{T}_{l_i} \neq \mathbf{T}_o$. Thus, if an optimization procedure gets trapped into $\mathbf{T}_{l_i} \neq \mathbf{T}_o$, a relatively small offset added on the obtained transformation can pull it out of $\mathcal{N}_{\mathbf{T}_{l_i}}$, and then a further chance for a following up optimization to converge to $\mathbf{T}_o$ is given. On the other hand, when an optimization procedure achieves $\mathbf{T}_o$, the same offset will not pull it out of $\mathcal{N}_{\mathbf{T}_o}$, and thus a following up optimization process still achieves $\mathbf{T}_o$. Practically, when an optimization procedure converges, in order to pull it out of $\mathcal{N}_{\mathbf{T}_{l_i}}$ ($\mathbf{T}_{l_i} \neq \mathbf{T}_o$), or to validate whether it reaches $\mathbf{T}_o$ or not, we propose to randomly add an reasonable disturbance (e.g. larger than $R_{\mathbf{T}_{l_i}}$ and much less than $R_{\mathbf{T}_o}$) on the obtained transformation and resume the optimization process. (For the determination of the range of disturbance, please refer to Section 3.) As a consequence, the potentiality of obtaining a robust optimization result increases. This process iterates until the change of obtained transformations at two successive iterations becomes sufficiently small.

### 2.3 Rotation Representation

The technique proposed in Section 2.2 is demonstrated on 3D rigid registration. We use quaternion algebra to represent 3D rotations in rigid transformations.

This is because any rotation can be uniquely represented by quaternion as a complex number:

$$\begin{aligned} q &= a + ib + jc + kd \\ &= \cos(\frac{\alpha}{2}) + i(x\sin(\frac{\alpha}{2})) + j(y\sin(\frac{\alpha}{2})) + k(z\sin(\frac{\alpha}{2})), \end{aligned} \tag{2}$$

where $(x, y, z)^T$ is an unit vector representing the axis of rotation, $\alpha$ is the angle of rotation. Note that other rotation representations may have ambiguity in some cases, e.g. Euler angle representation has ambiguity when the attitude (or elevation) is equal to $\pm\frac{\pi}{2}$. Moreover, quaternion has another advantage of encoding rotation in an order-independent manner.

In addition, in Equation 2, we have $a^2 + b^2 + c^2 + d^2 = 1$. In order to decrease the degree of freedom of quaternion for optimization, Equation 2 is revised as follows,

$$\begin{aligned} q &= 1 + i\tilde{b} + j\tilde{c} + k\tilde{d} = 1 + i\frac{b}{a} + j\frac{c}{a} + k\frac{d}{a} \\ &= 1 + i(x\tan(\frac{\alpha}{2})) + j(y\tan(\frac{\alpha}{2})) + k(z\tan(\frac{\alpha}{2})). \end{aligned}$$

Furthermore, during optimizations of brain image registrations, large rotations from optimal position, e.g. $\geq \frac{\pi}{2}$ and $\leq -\frac{\pi}{2}$, are fatal and most likely to invert brains in floating images. To avoid this, a hard constraint, namely $\tilde{b}^2 + \tilde{c}^2 + \tilde{d}^2 < 1$, may be added to limit the search space of rotation angle to $\alpha \in (-\frac{\pi}{2}, \frac{\pi}{2})$.

## 3 Implementation Details

In our implementation, in order to accelerate the registration process, a multi-resolution approach based on the Gaussian Pyramid representation [1, 14, 2] is exploited. Four resolution levels are used and the definition of resolution levels in the Gaussian Pyramid representation follows the same convention as in [1], i.e. Level 0 image represents the highest and original resolution and Level 3 image represents the lowest resolution. Smoothing is performed via the binomial filter with coefficients $[1, 4, 6, 4, 1]$ [14]. For the ease of implementation, all voxels in the downsampled floating volumes are used at Levels 1 – 3. At Level 0, 1/4 (25%) of all voxels are sampled (one voxel randomly picked from every $2 \times 2$ matrix in each slice). To construct the joint and marginal histograms, the image intensity values are linearly scaled to 64 bins, which have been commonly used in the mutual information based registrations. For optimization at each resolution, we use the Powell's direction set method [12] with Brent's 1D line minimization, where the fractional precision convergence parameters are set to $10^{-4}$ and $10^{-3}$ respectively.

The transformation disturbance across progressive optimizations as discussed in Section 2.2 is only performed at Level 3 for reasons of speed, since the computational burden at higher levels can be more significant. On the other hand,

it is empirically observed that, although from large initial misalignments, the majority of optimization results can converge close to the ground truth at Level 3, with the help of transformation disturbance. (Please refer to Section 4.2 for details.) For an optimization at Level 3, a disturbance, which consists of 3 disturbances for translational parameters and 3 disturbances for quaternion parameters, is randomly generated based on Gaussian distributions with zero means. To determine the standard deviations of these Gaussian distributions, a set of experiments with different standard deviation values have been performed. Empirical results suggest that a relative high success rate is given by the following set of standard deviation values: for translation, they are equal to 1/16 of image dimension sizes (in millimeter), and for quaternion, they are equal to $\tan(\pi/16)$. The convergence criterion is that, at two successive iterations, the change of obtained transformations for an individual degree of freedom is less than 1/5 of the corresponding standard deviation.

## 4 Experimental Results

### 4.1 Image Datasets and Ground Truth

In the experiments described below, we used a set of real CT – T1 data obtained from the Retrospective Image Registration Evaluation (RIRE) project[1]. Note that all the T1 images have been rectified for intensity inhomogeneity and scaling. In general, the size of a CT image volume is $512 \times 512 \times 34$ voxels and the voxel size is $0.65 \times 0.65 \times 4mm^3$, and a T1 image contains $256 \times 256 \times 26$ voxels of dimensions $1.25 \times 1.25 \times 4mm^3$.

With regard to the data, we determined the "ground truth" for registration experiments as follows. First, the multi-resolution mutual information based and normalized mutual information (NMI) [13] based methods were used to register the image pairs. The evaluations of accuracy were obtained from the RIRE project. By examining the median errors, four datasets (Datasets pt-001, pt-003, pt-005 and pt-007) with less than 1mm registration error were selected and used in the experiments. Then the corresponding optimal transformations, whose median errors were 0.5077 (for pt-001), 0.7200 (for pt-003), 0.7807 (for pt-005) and 0.6179 (for pt-007) respectively, were used as the ground truth registrations.

### 4.2 Justification and Determination of Disturbance

In order to justify the capability of the proposed transformation disturbance on helping optimization to avoid local maxima, especially at Level 3, randomized experiments were performed for the mutual information based registration method with disturbance and quaternion (hereafter referred to as MI-d), and also for the conventional mutual information based registration method (hereafter referred

[1] *Images were provided as part of the project, "Evaluation of Retrospective Image Registration", National Institutes of Health, Project Number 8R01EB002124-03, Principle Investigator, J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN.*

to as MI) as a comparison. The testing image pair was the aforementioned pt-001 CT – T1 dataset and the experiment for either method took 100 trials. At each trial, the ground truth registration parameters of pt-001 were perturbed by six uniformly distributed random offsets for all translational and rotational axes. The perturbed parameters were then treated as the starting alignment. In order to show the high optimization capability of MI-d with respect to initial alignment, random offsets for X and Y translational axes were drawn between around [-150, 150]mm and those for Z translational axis were drawn between around [-70, 70]mm. (Note that these ranges were set so that two brains in CT and T1 images have at least 10% overlapping region.) While, random offsets for each rotational axis were respectively drawn between [-0.52, 0.52] radians, i.e. [-30, 30] degrees. (The perturbed rotations were first converted to quaternions prior to MI-d.) As a fair comparison, the same set of randomized starting alignments was used for MI and MI-d.

In addition, as mentioned in Section 3, in order to determine a suitable setting for generating random disturbances, we have tested several different sets of standard deviation values for the Gaussian distributions: $1/n$ of image dimension sizes (in millimeter) for translation, and $\tan(\pi/n)$ for quaternion, $n \in \{8, 16, 32, 64\}$. Then, we selected the one set, which produced the most *successful* results, for further experiments. Since the registrations were only performed to the downsampled image pair at Level 3, an optimization result was judged to be *successful* when the individual translational errors w.r.t. the ground truth were less than 10mm and the individual rotational errors (measured by Euler angles) were less than 5°. Such thresholds were selected because, based on our empirical experiences, a starting alignment within them definitely can be fine-tuned to converge to the ground truth.

Table 1 presents the number of successful optimization results for MI and MI-d with different sets of standard deviations (i.e. MI-d-8, MI-d-16, MI-d-32, MI-d-64) on pt-001 dataset at Level 3. In the "# Success" column, it is suggested that, as compared with MI, MI-d with different sets of standard deviations can give much more successful optimization results at Level 3. It is also noted that, amongst the four different sets, MI-d-16 has the best performance and thus we determine to adopt it for further experiments. The average running time at each trail for each method is also listed in the table. Obviously, all MI-d methods take

**Table 1.** The number of successful optimization results and the average processing time (in seconds) of MI, MI-d with different sets of standard deviations on pt-001 CT – T1 dataset at Level 3

| | # success | Time (sec) |
|---|---|---|
| MI | 47 | 28 |
| MI-d-8 | 90 | 227 |
| **MI-d-16** | **96** | **193** |
| MI-d-32 | 83 | 311 |
| MI-d-64 | 81 | 321 |

relatively longer time to converge than MI. However, as we do not perform transformation disturbance at higher levels (i.e. Levels $0-2$), where the computational burden of conventional MI is already significant, the difference between the overall processing time of MI-d and MI will be much less than that at Level 3.

### 4.3 Performance Comparisons on Robustness and Accuracy

In this section, we further study and compare the registration performance (w.r.t. robustness and accuracy) of MI-d-16 and MI. A series of similar randomized experiments as described in Section 4.2 was performed on the aforementioned four clinical CT – T1 image pairs (i.e. pt-001, pt-003, pt-005 and pt-007) with all four resolution levels. Therefore, for either method, an experiment with 100 trials was performed on each dataset with a set of randomized starting alignments generated as described in Section 4.2.

To evaluate each derived registration with respect to the corresponding ground truth registration, similar to [8, 6], a tight bounding box was fitted around the brain for each T1 images. For each of the eight corner points, the Euclidean distance between the ground truth position and the position transformed by our solution was computed. The median value of the eight distances was then taken for assessing registration success. A registration was judged to be successful if the median error was smaller than or equal to 4mm, which was the largest voxel dimension of the CT – T1 image pair; otherwise, it was considered a misregistration.

Table 2 lists the success rates for MI and MI-d-16 for all testing image pairs (pt-001, pt-003, pt-005 and pt-007), together with the means and standard deviations of the median errors (in millimeters) for the *successful* registrations. It is shown in the table that MI-d-16 consistently gives much higher success rates as compared with MI. For registration accuracy, it is observed that the median errors of the successful registrations for MI-d-16 are comparable to those for MI, and are acceptably low.

**Table 2.** The success rates of MI and MI-d-16, and the means and standard deviations of the median errors (in millimeters) for different testing image pairs

| | MI | | MI-d-16 | |
|---|---|---|---|---|
| | success% | mean ± sd | success% | mean ± sd |
| pt-001 | 51% | 0.5295 ± 0.0459 | 96% | 0.5438 ± 0.0530 |
| pt-003 | 45% | 0.1983 ± 0.1103 | 99% | 0.2134 ± 0.2854 |
| pt-005 | 40% | 0.5182 ± 0.7830 | 93% | 0.3959 ± 0.6452 |
| pt-007 | 44% | 1.6842 ± 0.4790 | 92% | 1.6959 ± 0.4138 |

## 5 Conclusion and Discussion

To conclude, this paper has proposed a simple and computationally efficient technique based on transformation disturbance to assist an optimization method

to avoid local optima, and hence to achieve a robust optimization result. The new technique has been demonstrated on 3D rigid registrations, and quaternions have been adopted to represent rotations for the purpose of the unique and order-independent expression. Randomized registration experiments on four clinical CT and MR-T1 datasets have revealed that the success rates of our method are consistently much higher than those of the conventional multi-resolution mutual information based method. It has been also shown that the registration accuracy of the new method is acceptably high.

Finally, we would like to note that there is no practical guarantee for our method to achieve global maximum. However, empirical observations have shown that the optimization performance of our method is much better than the conventional multi-resolution approach. A theoretical justification of the work is desirable and remains a topic for future research. Furthermore, although the illustration and demonstration in this paper just concentrate on mutual information as similarity measure and the Powell's optimization method, the proposed technique is quite general and can be applied to other similarity measures (e.g. normalized mutual information and so on) and optimization methods (e.g. simplex methods and so on).

## Acknowledgements

The authors would like to acknowledge the support from the Research Grants Council (RGC) of Hong Kong under Grant HKUST6155/03E.

## References

1. P.J. Burt and E.H. Adelson. The Laplacian Pyramid as a Compact Image Code. *IEEE Trans. Comms.*, 31(4):532–540, 1983.
2. A.A. Cole-Rhodes, K.L. Johnson, and et al. Multiresolution Registration of Remote Sensing Imagery by Optimization of Mutual Information Using a Stochastic Gradient. *IEEE Trans. Image Processing*, 12(12):1495–1511, 2003.
3. T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.
4. S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721 – 741, 1984.
5. M. Jenkinson and S. Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143 – 156, 2001.
6. Z.F. Knops, J.B.A. Maintz, M.A. Viergever, and J.P.W. Pluim. Normalized mutual information based registration using k-means clustering and shading correction. *Medical Image Analysis*, 2005 (To appear).
7. S. Kullback. *Information Theory and Statistics*. Dover Publications, Inc., 1968.
8. F. Maes, A. Collignon, and et al. Multimodality Image Registration by Maximization of Mutual Information. *IEEE Trans. Med. Img.*, 16(2):187–198, 1997.
9. G.P. Penney, J. Weese, and et al. A Comparison of Similarity Measures for Use in 2D-3D Medical Image Registration. *IEEE Trans. Med. Img.*, 17(4):586–595, 1998.

10. D. Plattard, M. Soret, J. Troccaz, P. Vassal, J.-Y. Giraud, G. Champleboux, X. Artignan, and M. Bolla. Patient set-up using portal images: 2d/2d image registration using mutual information. *Comput Aided Surgery*, 5(4):246 – 262, 2000.
11. J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Image Registration by Maximization of Combined Mutual Information and Gradient Information. *IEEE Trans. Med. Img.*, 19(8):809–814, 2000.
12. W.H. Press, S.A. Teukolsky, and et al. *Numerical Recipes in C, 2nd Edition*. Cambridge University Press, 1992.
13. 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.
14. W.M. Wells, P. Viola, and et al. Multi-Modal Volume Registration by Maximization of Mutual Information. *Medical Image Analysis*, 1(1):35–51, 1996.

# MR-CT Image Registration in Liver Cancer Treatment with an Open Configuration MR Scanner

Songyuan Tang[1,2], Yen-wei Chen[1], Rui Xu[1], Yongtian Wang[2], Shigehiro Morikawa[3], and Yoshimasa Kurumi[3]

[1] College of Information Science and Engineering, Ritsumeikan University, Japan
sytang@bit.edu.cn, chen@is.ritsumei.ac.jp
gr042049@se.ritsumei.ac.jp
[2] Department of Opto-electronic Engineering, Beijing Institute of Technology, P.R. China
wyt@bit.edu.cn
[3] Shiga University of Medical Science, Japan
morikawa@belle.shiga-med.ac.jp

**Abstract.** MR – CT image registration has been used in the liver cancer treatment with an open MR Scanner to guide percutaneous puncture for ablation of tumors. Due to low magnetic field and limited acquisition time, MR images do not always show the target clearly. Sometimes, assistance of CT images is helpful for the navigation to the target. The shape of the liver within the surgical procedure is different from that of preoperative CT images due to the patient position for the convenience of surgery. It is quite difficult to match the images accuracy during surgery. In this paper, we have proposed a method to improve the registration accuracy of images with an open MR scanner and preoperative CT images of the liver. The method includes three parts. Firstly a semiautomatic method is used to extract the liver from MR and CT images as region of interest (ROI). Then, an affine registration is used to match the images roughly. Finally, BSpline-based nonrigid registration is applied. The results are found to be satisfactory with visual inspection by experts and with evaluation by the distance of two liver surfaces, while comparing with other methods.

## 1 Introduction

Recently, various minimally invasive treatments have been widely spread for the treatment of liver tumors. For the image guidance, ultrasonography has been mainly used. With developments of open configuration MR scanners, MR images have also been utilized for the navigation of minimally invasive therapies, because MR images have many advantages for image navigation, such as good soft tissue contrast, free from ionizing radiation and multiplanar capabilities. Microwave ablation, an established procedure for the treatment of liver tumors, has been successfully combined with MR image guidance [1]. At the initial stage of surgery, only real-time MR image, which is continuously acquired with gradient echo sequence within 2 seconds, is used for image guidance. The image planes including the path of the needle are interactively controlled by surgeons using an optical tracking system. The contrast of real-time MR images acquired within 2 seconds is not always satisfactory. The cases, in which real-time MR images can show the target clearly, are selected for this treatment. As the second stage,

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 289 – 296, 2006.

navigation software on an external computer is used in combination with real-time MR images [2, 3]. After the patient position for the treatment is fixed, multi-slice MR images are acquired and transferred to the software as 3D volume data. Real-time MR images in the corresponding planes are simultaneously displayed on the surgeon's monitor. Multi-slice MR images just before the treatment are much better in the image quality than real-time MR images. In addition, the tumor region is manually traced and shown with color on the display. The assistance of this software is quite useful and remarkably expands the indication of this treatment. In some cases with severe chirrhosis, however, visualization of the tumor is still difficult with MR images acquired by the open configuration MR system. In such cases, the combination of preoperative CT images will be greatly helpful, if CT images can be registered to the position of MR images accurately.

Liver image registration goes back to 1983 [4], but most works on liver registration are done in recent years [5]-[9]. Only one paper is found about MR liver image registration with open MR system [9]. In this paper, intensity-based and rigid transform registration is used. There are no previous reports of liver registration using intensity-based nonrigid registration. Since nonrigid registration can deform an image, it is possible to get better result than that of rigid registration.

In this paper, we have developed a non-rigid registration method to match MR and CT images in the liver tumor treatment with open MR system. Since CT images are acquired before surgery, we have enough time to carefully segment it by manual and obtain accuracy CT images of the liver. When multi-slice MR images are acquired during surgery, deformable model is adopted first to roughly segment the images, and then the roughly segmented images are trimmed manually. After segmentation, the intensity-based affine registration is applied to both segmented images, and then BSpline-based nonrigid registration is applied. Both registrations use mutual information (MI) as similarity metrics since MI have been proven robust in the multi-model image registration [10]. The registration accuracy is accessed by the visual inspection and the distance of two liver surfaces. Compared with the Andres' method [9], the results are satisfied.

## 2 Method

### 2.1 Liver Segmentation of MR Images

Since the intensity of liver and other tissues in abdomen is similar in the MR images, segmentation of the liver is difficult. Both consideration of speed and accuracy, we combine the automatic and the manual method to segment the liver. Firstly, we adopt automatic segment to segment liver roughly and remove the most non-liver part, such as wall of abdomen, spine etc, and then manually segment liver organ with a little effort. We adopt the deformable surface model to segment liver due to its robust and fast. The liver surface and soft tissue are modeled by a surface tessellation using connect triangles. The initial model is a tessellated sphere, which is put at the center of the image. Each vertex of the sphere surface is updated to approach the liver or soft tissue surface. When the vertex of the sphere surface

approaches the liver or soft tissue surface enough, the updating is stopped. Detailed can be found in refs. [11][12].

### 2.2 Affine Registration

Mutual information is the most robust similarity metrics for the multimodal image registration and is widely used [13][14]. Therefore, we select it to match the two liver images. We use affine registration firstly to match two images roughly. The affine transformation is $T=\{t_x,t_y,t_z,r_x,r_y,r_z,s_x,s_y,s_z,\gamma_{xy},\gamma_{yz},\gamma_{zx}\}$, where $t_x,t_y,t_z$ are translations along the $x,y$ ,and $z$ axes respectively, $r_x,r_y,r_z$ are rotations about the $x,y$ ,and $z$ axes respectively, $s_x,s_y,s_z$ are scales about the $x,y$ ,and $z$ axes respectively, and $\gamma_{xy},\gamma_{yz},\gamma_{zx}$ are shears about $x-y,y-z$ and $z-x$ plane respectively.

The initial translations are determined by the centers of mass of the two segmented livers, the initial scales are set to one, and the initial rotations and shears are set to zero. The Powell optimization [15] is used to obtain these optimal parameters.

### 2.3 BSpline-Based Nonrigid Registration

BSpline-based free form deformation (FFD), which is firstly proposed [16] to process breast images, is used to deform the CT image. The shape of image space can be controlled by changing the control grids of the BSpline, and the transform is Cn-1 smooth continuous, where n is the order of BSpline basis function. Usually the C2 continue is enough, therefore, we select three order of BSpline basis function to deform the image space as followed.

$$\begin{aligned}\theta_0(s)&=(1-s)^3/6\\ \theta_1(s)&=\left(3s^3-6s^2+4\right)/6\\ \theta_2(s)&=\left(-3s^3+3s^2+3s+1\right)/6\\ \theta_3(s)&=s^3/6\end{aligned}\tag{1}$$

The deformation field defined by FFD can be represented as:

$$\vec{u}(x,y,z)=\sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3}\theta_l(u)\theta_m(v)\theta_n(w)P_{i+l,j+m,k+n}\tag{2}$$

Here $P$ is the control grid.

Usually the cost function includes two parts, one is similarity metrics, which characterizes the similarity of two images, and the other is deformation, which is associated with the particular deformations [17].

$$C=-C_{similarity}+C_{deformation}\tag{3}$$

But computing cost increases greatly when the deformation part is added. Considering the C2 continuous of BSpline, when the two images aren't different greatly, we can delete the deformation part and still get good results. Actually, after affine transform, the two images are very similar. Therefore we only use similarity part in the cost function so as to save much time.

## 3 Experiment Results

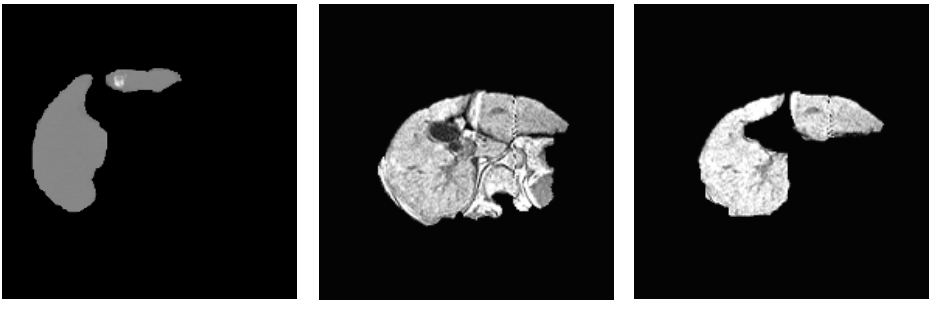
### 3.1 Data Acquisition

CT images were acquired with Somatom Sensation Cardiac/16 (Siemens). The size of CT images is $512 \times 512 \times 25$, slice thickness is 7.0 mm, and the in-plane dimensions are $0.58mm \times 0.58mm$. CT images were acquired 3 days before the microwave ablation. MR images were acquired by a 0.5 T open configuration MR system, Signa SP/I (GE Healthcare). The size is $256 \times 256 \times 28$, slice thickness is 5.0 mm, and the in-plane dimensions are $1.172mm \times 1.172mm$. In this case, laparoscopic guidance was combined with MR image guidance. The abdominal cavity was inflated with $CO_2$ gas for the laparoscopy. MR images were acquired after the inflation during surgical procedure.

### 3.2 Liver Segmentation

An example of manual segmentation of CT is shown in figure 1(a). It is segmented by experts carefully.

The deformable surface model segments the Open MRI very quickly, it's about 1 second. An example is shown in figure 1 (b). Then manual segmentation is used to trim the result further. The segmentation result is shown in figure 1 (c). It's need about 10-15 minutes.



(a) (b) (c)

**Fig. 1.** (a) An example of manual segmentation of CTA (b) deformable surface model segmentation of Open MRI (c) manual segmentation of the result of (b)

### 3.3 Results of Registration

After the livers are segmented, the images sizes can be deduced greatly. We reduce the size of CT image to $340 \times 340 \times 25$ , and the size of MR image to $180 \times 180 \times 28$ .

An example is shown in figure 2 and 3. Fig.2 (a) is a slice of CT image and Fig.2 (b) a slice of Open MR image before registration. Fig.2 (c)-(e) is the results of CT images after rigid, affine and BSpline registration respectively, and Fig.2 (f)-(h) is the tumor extracted from registered CT images overlap on the MR image respectively.

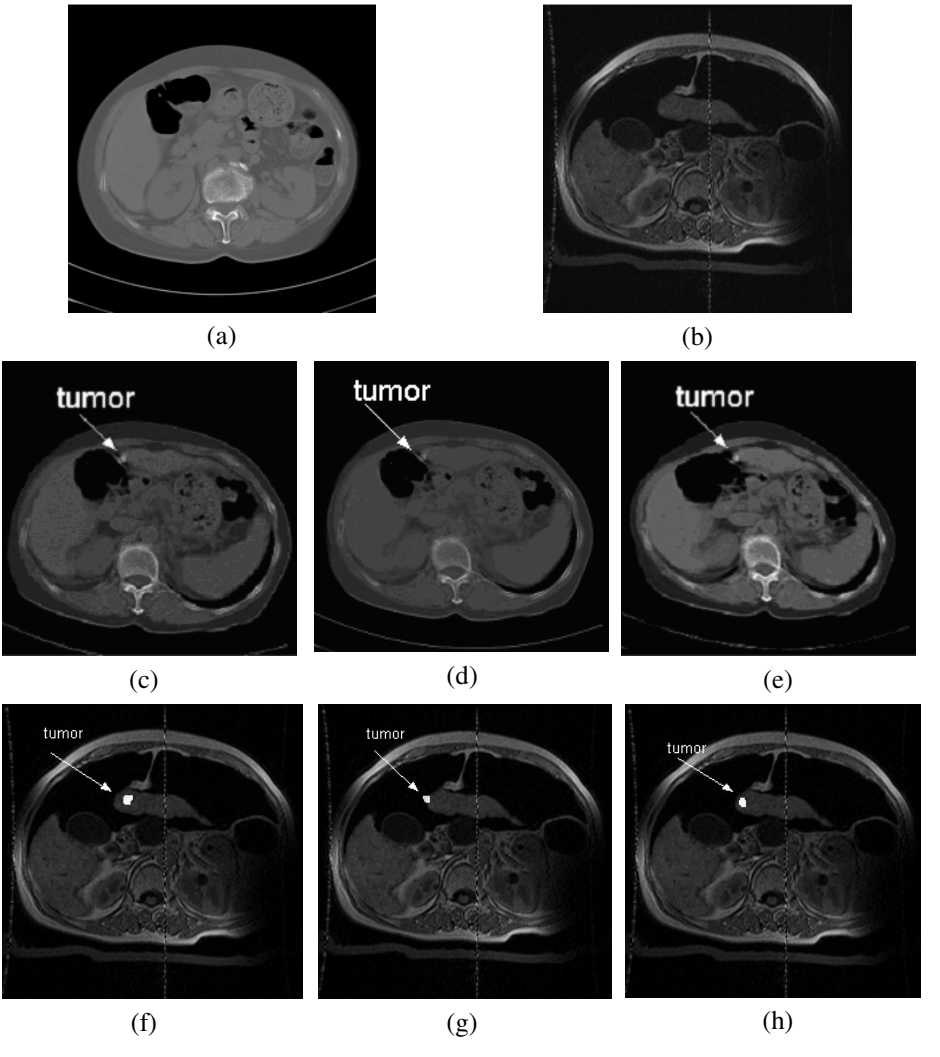
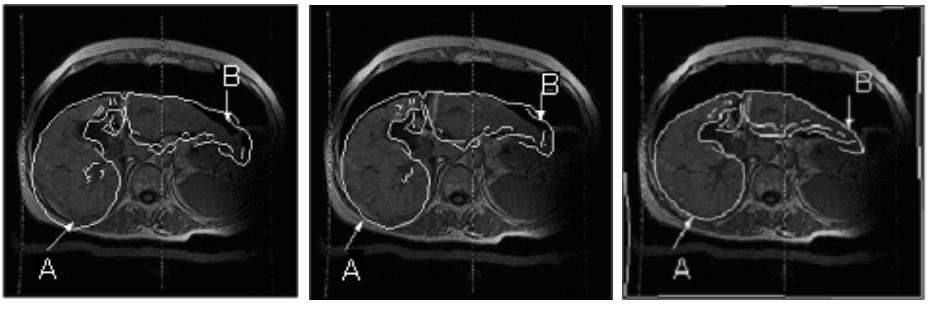



**Fig. 2.** (a) a slice of CTA (b) a slice of Open MRI (c)-(e) results of rigid registration, affine registration , and BSpline registration (f)-(h) tumor extracted from CTA after rigid registration, affine registration , and BSpline registration and overlap on the Open MRI

The contour of CT images after rigid, affine registration and BSpline-based nonrigid registration are overlapped on the Open MRI, and an example is shown in figure 3.

The computing costs are listed in table 1.




(a) (b) (c)

**Fig. 3.** The contour of CTA overlapped on the Open MRI (a) rigid registration (b) affine registration (c) BSpline-based nonrigid registration

**Table 1.** Computing cost

| | Proposed method (second) | Andres' |
|---|---|---|
| Segment liver from Open MRI | 900 | |
| Affine registration | 500 | |
| Nonrigid registration | 200 | |
| total | 1600 | 990 |

### 3.4 Algorithm Evaluation

We evaluate our method from visual inspection, distance of liver surfaces, and computing cost, and compare the results of the proposed method with those of Andres'.

Andre's method is rigid registration, only three translations and three rotations are used to match images. An example is shown as follow.

**Visual inspection.** Figure 2 (c), (d), and (e) are the results of CT images after rigid, affine and BSpline registration. We can find the tumor is just beneath the surface of the liver. Figure 2 (f)-(h) show the tumor overlapped on the MR images during surgery. The tumor is extracted from the results of rigid, affine and BSpline registration respectively. It is easy to find that the position of the tumor is far away from the liver surface after rigid registration, a part of the tumor is outside theliver after affine registration, and the tumor position is just beneath the liver surface after BSpline registration. Therefore after BSpline registration, we can get satisfied result.

In figure 3, from positions pointed by arrow A and B, we can easily find the liver surface of CT images after BSpline registration approaches the liver surface of MR images best, that of affine registration is better than that of rigid registration.

From the visual inspection, the accuracy of proposed method is much better than Andres'.

**Distance of liver surface.** We use the method mentioned in reference [9] to estimate the distance of two liver surfaces. Detailed can be found in the reference. The 3D error of the proposed method is about 1.5 mm, while Andre's about 3mm.

**Computing Cost.** The cost is listed in table 1. Since CT images are acquired before surgery, there is enough time to segment liver from it. We don't include this time in the proposed method. We only consider the time in the surgery. Using deformable model and manually segmentation, it costs about 900 seconds. Affine registration needs about 500 seconds, and BSpline-based nonrigid registration about 200 seconds. The total is about 1600 second that is about 27 minutes, while Andres's only needs about 900 seconds.

## 4 Discussion and Conclusion

This case was extraordinary one, because MR image and laparoscopy were combined for the image navigation. The inflation of the abdominal cavity with $CO_2$ gas caused bigger deformation of the liver than usual cases. Even in such a special case, the proposed method can be successfully applied to CT and MR image registration. Although the computing time is more than Andres', 27 minutes can be accepted. The high accuracy of the proposed method can be more effective to remove the tumor of patients. Actually, the most time is used for segmentation of Open MR in the proposed method. If the automatically segmentation algorithm is more effective, the time will reduce greatly. In the future, we'll research better liver segmentation method to reduce the computing cost needed in the proposed method.

## Acknowledgement

This work was supported partly by the Strategic Information and Communications R&D Promotion Programme (SCOPE) under the Grand No. 051307017 and the National Key Basic Research and Development Program (973) Grant No. 2003CB716105.

## References

1. Morikawa S, Inubushi T, Kurumi Y, Naka S, Sato K, Tani T, Yamamoto I, Fujimura M. MR-Guided microwave thermocoagulation therapy of liver tumors: initial clinical experiences using a 0.5 T open MR system. J Magn Reson Imaging 16: 576-583, 2002.
2. Morikawa S, Inubushi T, Kurumi Y, Naka S, Sato K, Demura K, Tani T, Haque HA, Tokuda J, Hata N. Advanced computer assistance for magnetic resonance-guided microwave thermocoagulation of liver tumors. Acad Radiol 10: 1442-1449, 2003.
3. Sato K, Morikawa S, Inubushi T, Kurumi Y, Naka S, Haque HA, Demura K, Tnai T. Alternate biplanar MR navigation for microwave ablation of liver tumors. Magn Reson Med Sci 4: 89-94, 2005.

4. A. Venot, , J. L. Golmard, J. F. Lebruchec, L. Pronzato, , E. Walter, G. Frij, and J. C. Roucayrol: Digital methods for change detection in medical images. In Deconinck, F. (ed.), Information processing in medical imaging, pp. 1–16, Dordrecht. Nijhoff publ, (1983).
5. C. K . Hoh, M. Dahlbom, G. Harris, Y. Choi, R. A. Hawkins, M. E. Phelps, and J. Maddahi: Automated iterative three-dimensional registration of positron emission tomography images. Journal of nuclear medicine, 34, (1993) 2009–2018.
6. A. M. Scott, H. Macapinlac, C. R. Divgi, J. J. Zhang, H. Kalaigian, K. Pentlow, S. Hilton, M. C. Graham, G. Sgouros, C. Pelizzari, G. Chen, J. Schlom, S. J. Goldsmith, and S. M. Larson: Clinical validation of SPECT and CT/MRI image registration in radiolabeled monoclonal antibody studies of colorectal carcinoma. Journal of nuclear medicine, 35, (1994) 1976– 1984.
7. A. M. Scott, H. Macapinlac, J. Zhang, F. Daghighian, N. Montemayor, H. Kalaigian, G. Sgouros, M. C. Graham, K. Kolbert, S. D. J. Yeh, E. Lai, S. J. Goldsmith and S. M. Larson: Image registration of SPECT and CT images using an external fiduciary band and three-dimensional surface fitting in metastatic thyroid cancer. Journal of nuclear medicine, 36, (1995) 100–103.
8. S. Mehrdad *et al*: Automatic CT–SPECT registration of livers treated with radioactive microspheres, Phys. Med. Biol. 49, (2004)
9. C. Andres, L.D. Jeffrey, S.L. Jonathan, and L.W. David: Semiautomatic 3-D Image Registration as Applied to Interventional MRI Liver Cancer Treatment, IEEE Transaction on Medical Imaging, Vol.19 (2000) 175–185
10. Josien Pluim: Information Based Registration of Medical Images, PhD thesis (2003)
11. A. Kelmen, G. Szekely, G. Gerig: Elastic Model-Based Segmentation of 3-D Neuroradiological Data Sets, IEEE Transaction on Medical Imaging, Vol.18, No. 10, (1999) 828 – 839
12. Stephen M. Smith: Fast Robust Automated Brain Extraction, Human Brain Mapping, Vol.17 (2002) 143–155
13. J.P.W. Pluim, J.B.A. Maintz, M.A. Viergever: f-Information measures in medical image registration, IEEE Transactions on Medical Imaging, vol. 23, (2004) 1506-1518.
14. W.M Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis: Multi-modelal volume registration by maximization of mutual information, Medical Image Analysis, Vol. 1, (1996) 35–51
15. H. William, A.T. Saul *et al*: http://www.library.cornell.edu/nr/bookcpdf.html
16. 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 Transaction on Medical Imaging, Vol. 18, No. 8, (1999) 712–721
17. J. V. Hajnal, D. L. G. Hill, D. J. Hawkes: Medical Image Registration, CRC Press (2001)

# Nonrigid Registration of Multitemporal CT and MR Images for Radiotherapy Treatment Planning

Pieter Slagmolen[1], Dirk Loeckx[1], Sarah Roels[2], Xavier Geets[3], Frederik Maes[1], Karin Haustermans[2], and Paul Suetens[1]

[1] Medical Image Computing (ESAT/PSI), Faculties of Medicine and Engineering
[2] Department of Radiation Oncology
[1,2] University Hospital Gasthuisberg, Herestraat 49, 3000 Leuven, Belgium
[3] Department of Radiation Oncology
Université Catholique de Louvain, St-Luc University Hospital, 10 Ave. Hippocrate, 1200 Brussels, Belgium
pieter.slagmolen@uz.kuleuven.ac.be

**Abstract.** External beam radiotherapy treats cancer lesions with ionizing radiation. Successful treatment requires a correct definition of the target volume. This is achieved using pre-treatment MR and CT images. However, due to changes in patient position, tumor size and organ location, adaptation of the treatment plan over the different treatment sessions might be wanted. This can be achieved with extra MR and CT images obtained during treatment. Bringing all images into a common reference frame, the initial segmentations can be propagated over time and the integrated dose can be correctly calculated.

In this article, we show in two patients with rectum cancer and one with neck cancer that a significant change in tumor position and shape occurs. Our results show that nonrigid registration can correctly detect these shape and position changes in MR images. Validation was performed using manual delineations. For delineations of the mandible, parotid and submandibular gland in the head-and-neck patient, the maximal centroid error decreases from 6 mm to 2 mm, while the minimal Dice similarity criterium (DSC) overlap measure increases from 0.70 to 0.84. In the rectal cancer patients, the maximal centroid error drops from 15 mm to 5 mm, while the minimal DSC rises from 0.22 to 0.57.

Similar experiments were performed on CT images. The validation here was infeasible due to significant inaccuracies in the manual delineations.

## 1 Introduction

External beam radiotherapy treats cancer lesions with ionizing radiation. Different radiation beam angles and shapes are used to obtain a maximal dose delivery to the target volume while minimizing the irradiation of the surrounding, healthy, tissue. Successful treatment and the application of new techniques like Intensity Modulated Radiotherapy (IMRT), require a correct definition of the target volume and surrounding tissues, a high radiation dose gradient and

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 297–305, 2006.

field resolution and an accurate, reproducible method for positioning the patient on the treatment table.

In most cases, a pre-treatment MR image is recorded to accurately delineate the tumor and surrounding organs of interest, while a pre-treatment CT image is required to calculate the dose distribution and design the treatment plan. Recently PET images, using one or more tracers, were introduced to provide additional functional information. The multimodal images of the same patient are brought into correspondence using rigid or affine image registration. However, except for the brain, the human body deforms nonrigidly. Therefore, the patient's position should be kept as constant as possible. This is achieved with e.g. personalized facial masks, laser-guided alignment to skin marks and portal images. These tools are used during the pre-operative scans as well as during the treatment.

However, geometrical uncertainties, such as internal organ movement, bladder and rectum distention, setup error, patient breathing and inevitable differences in imaging position can still cause a major change in the location of the target volume [1]. Also, shrinking of the tumor and possible weight loss of the patient oppose an accurate patient positioning and thus reduce the benefits of IMRT.

To optimize treatment, the total dose is distributed over different treatment sessions. Usually, the treatment plan is worked out on the planning CT, acquired before radiotherapy, and it is not updated during the treatment. The planning CT is also used to calculate the total dose delivery, integrated over the different sessions. Due to the geometrical uncertainties and changes in the patient, errors can occur in the total dose calculation and distribution.

When images recorded during treatment are included along with the pre-treatment images, the treatment can be optimized. However, this is only rarely performed as it requires time-consuming re-delineation of the tissues for every session. In addition, the problem of a correctly integrated dose calculation still persists. A method, using a registration based on the contoured volumes to bring the dose distributions in correspondence, has already been developed [2].

We propose a novel method, that uses voxel-based nonrigid registration to bring the images, recorded with various modalities and at different time points, into correspondence. Next, the tissue delineations from e.g. the initial MR image can be propagated to the other MR images. The method is validated by comparing the automatically propagated delineations to manual delineations.

First, an overview of the used materials and methods is presented. Next, the results are demonstrated on rectum and head-and-neck images. At the end, we give a short discussion to finish with the conclusion.

## 2 Methods and Materials

### 2.1 Images

The method was applied to images obtained from three patients: two patients with a rectum tumor (rectum patients) and one patient with a neck tumor (head and neck patient).

*Rectum.* Each of the rectum datasets consists of 3 PET/CT and 3 MR images. Before each scan, the patients were asked to empty their bladder. The images were recorded prior to (week 0), during (week 3) and after (week 10) radiotherapy. Each week, the PET/CT and MR images were recorded at approximately the same time.

The PET/CT images were recorded on a Siemens Biograph 2 PET/CT scanner. The CT images have a dimension of $512 \times 512 \times 94$ voxels and a voxelsize of $0.98 \times 0.98 \times 3.00$ mm. The FDG-PET images have the same slice distance as the CT images, but an in-plane resolution of $128 \times 128$ pixels and thus a voxelsize of $3.92 \times 3.92$ mm. Due to the recording setup, the PET images are a priori registered to the CT images. To mimic the treatment position, the patient is positioned on a bellyboard (prone). The CT taken at week 0 is used as the planning image. All CT images are taken after administration of intravenous contrast. For the CT at week 3, rectal contrast was administered.

The MR images (dimensions $512 \times 416 \times 40$ voxels with voxelsize $0.78 \times 0.78 \times 4.9$ mm) were recorded on a Siemens SonataVision with a T1 sequence. Since the bellyboard is not suited for use with the MR scanner and since the prone position causes strong breathing artifacts in MR, the MR images are recorded with the patient in the supine position.

*Head-and-neck.* The third set of images consists of two CT and two MR images of a patient with a neck tumor, recorded prior to (week 0) and during (week 4) treatment. Again, the patient was recorded on both modalities at approximately the same time. To minimize pose differences, a personalized facial mask was used for the positioning of the patient.

The CT images were recorded on a Philips Mx8000 IDT 16 scanner, have a voxelsize of about $0.5 \times 0.5 \times 1$ mm and a dimension of $512 \times 512 \times 228$ voxels. For the recording of the MR images, a Philips Gyroscan Intera was used. The MR images have a voxelsize of about $0.5 \times 0.5 \times 3.5$ mm and a dimension of $512 \times 512 \times 50$ voxels.

A trained expert delineated the tumor in all rectum images and the parotid, mandible and submandibular gland in all head-and-neck images. The delineations were transformed from curves specified in each slice to a three-dimensional binary image.

### 2.2 Nonrigid Image Registration

Nonrigid registration is performed using a B-spline transformation model. A grid of mesh control points is positioned over the image and the displacements of these control points act as parameters for the deformation field. A gradual refinement of the grid allows more local deformations to be modeled.

Registration is driven by minimization of the cost function $E_c$, which consists of a similarity term $E_{\text{sim}}$ and two penalty terms $E_{\text{smt}}$ and $E_{\text{vol}}$. They are weighted with respectively weighting factors $\omega_{\text{sim}}$, $\omega_{\text{smt}}$ and $\omega_{\text{vol}}$:

$$E_c = \omega_{\text{sim}} E_{\text{sim}} + \omega_{\text{smt}} E_{\text{smt}} + \omega_{\text{vol}} E_{\text{vol}}. \tag{1}$$

The similarity term models the match between both images. We use mutual information (MI) of corresponding voxel intensities [3, 4] as the similarity measure. MI is based on the statistical relationship between both images to be registered, and has proven to be very robust for inter- and intramodal registration. A Parzen window approach is used to construct the joint histogram, as this improves the smoothness of the similarity function and allows for the calculation of analytical derivatives [5].

The penalty terms $E_{\text{smt}}$ and $E_{\text{vol}}$ will disfavor impossible or unlikely transformations, promoting respectively a smooth transformation field and local volume conservation.

Optimization is carried out using a multiresolution approach. Starting from downscaled images and a coarse mesh, the image and/or mesh resolution are increased at each stage. Within each stage, the optimal set of parameters is sought using a limited memory quasi Newton optimizer [5].

### 2.3 Validation

The transformation obtained from the nonrigid registration algorithm was applied to the binary delineation images. Thus, in the reference image we dispose of the original, reference delineation and different delineations obtained from each other co-registered image.

The registration was validated using three different criteria: distance between the centroids of corresponding delineations, their volume difference and volume overlap. Volume overlap is measured using Dice similarity criterium [6] defined as:

$$DSC(S_1, S_2) = 2\frac{V(S_1 \cap S_2)}{V(S_1) + V(S_2)} \tag{2}$$

where $S_1$ and $S_2$ are the two delineations.

If both volumes are identical, the DSC has a value of 1. When there is no overlap it is zero. A value of 0.7 indicates a great level of coincidence between segmentations. [7, 8].

## 3 Experiments and Results

### 3.1 Segmentation

An overview of tumor volume, as obtained from the manual segmentation over the different images, is given in Table 1. As expected, the tumor volume decreases during treatment. However, a difference in tumor volume of up to 30% between CT and MR delineations at the same time also occurs. As most tissues in the human body can be considered as incompressible, this was not expected. It is probably caused by inferior soft-tissue properties of CT images. The experts who performed the delineations confirmed that they had significantly more difficulties to correctly assess the tumor in the CT images than in the MR images.

### 3.2 Registration

The ultimate goal of our work is to put the different images of one patient in the same reference frame, preferably the planning CT.

Several scenarios exist to achieve this goal, including the registration of the CT and MR images at the same time point and the multitemporal registration of the CT and MR images recorded in between the different treatment sessions.

Because of the difference in image position and its multimodal character, the registration between MR and CT is much more difficult than the registration between different MR images. Especially in the rectum patients, due to the prone/supine positioning, very large deformations can be observed with tumor displacements of up to 7 cm.

In the images recorded at different time points, actual changes in the patients occur. Not only could they have gained or lost weight but, due to the radiation therapy, the tumor volume itself shrunk significantly (Table 1). For these reasons a severe volume penalty is not feasible.

**Table 1.** Volume of rectum tumors (a) and head and neck structures (b) during the course of radiotherapy. The volume is given in ml and is calculated based on the delineations on the MR images and the CT images.

| | | week 0 | week 4 | week 11 |
|---|---|---|---|---|
| patient 1 | MR | 25.7 | 11.5 | 8.7 |
| | CT | 18.41 | 9.58 | 4.98 |
| patient 2 | MR | 14.1 | 13.4 | 11.4 |
| | CT | 17.90 | 11.28 | 7.48 |

(a)

| | | week 0 | week 4 |
|---|---|---|---|
| mandible | MR | 53.41 | 49.13 |
| | CT | 66.98 | 66.31 |
| mandibular gland | MR | 25.99 | 16.60 |
| | CT | 26.64 | 14.38 |
| parotid | MR | 8.38 | 6.35 |
| | CT | 9.51 | 5.94 |

(b)

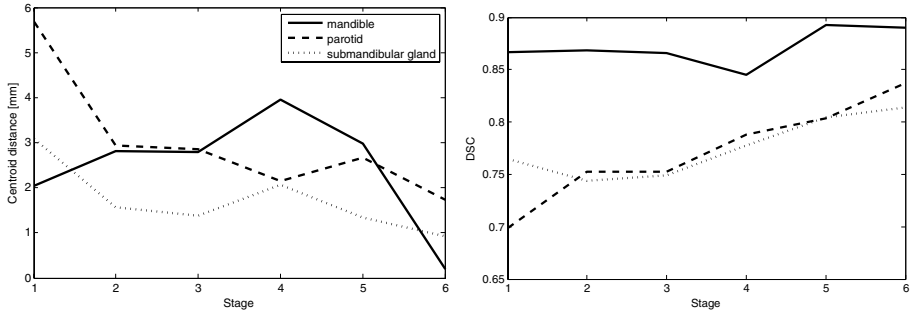



**Fig. 1.** Evolution of the centroid distance and the Dice similarity criterium over the course of our multiresolution approach for the head and neck patients. The centroid distance globally decreases over the different stages, although some increases occur. The DSC evolves more smoothly.

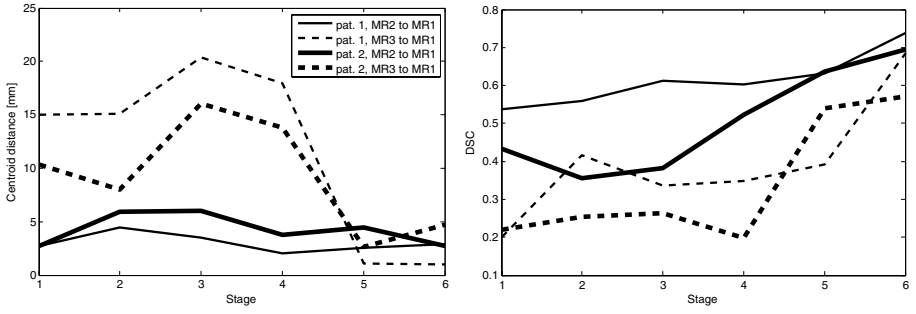
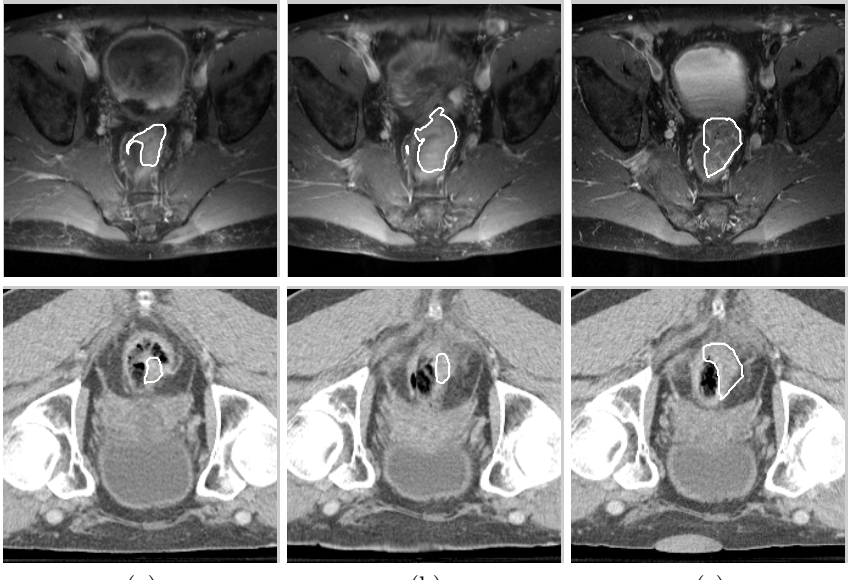



**Fig. 2.** Evolution of the centroid distance and the Dice similarity criterium over the course of our multiresolution approach for the rectum patients. The overall improvement is much better than the head-and-neck patient, but the final result is still not as good in terms of centroid distance and DSC.



(a) (b) (c)

**Fig. 3.** Visual assessment of the registration quality for MR-MR and CT-CT registration. (a) are the images at week 10, (c) are the images at week 0 and (b) is the registration of (a) to (c). The contours on (a) and (c) were manually delineated. The contour on (b) was deformed from (a) to (c). For the MR-MR registration, DSC improved from 0.54 between (a) and (c) to 0.81 between (b) and (c). For the CT-CT registration, although visually the registration seems good, the DSC did not improve because of a probable error on the delineation.

Figure 1 gives an overview of the registration results obtained for the head-and-neck images over the different stages, starting from affine transformation in stage

1. The centroid distance globally decreases over the different stages, although some increases occur. The DSC evolves more smoothly. The affine volume difference for the mandible, submandibular gland and parotid is respectively 4.27, 2.43 and 9.36 ml and decreased to 2.67, 1.59 and 4.16 ml after nonrigid registration.

The results obtained on the rectum patients are provided in Figure 2. The initial registration is much worse compared to Figure 1. While the overall improvement is much better than the head-and-neck patient, the final result is still not as good in terms of centroid distance and DSC. The affine volume difference for patient 1 decreases from 15.68 to 1.14 ml for registration with the MR at week 4 and from 16.90 to 4.41 ml for registration with the MR at week 11. Similar results are obtained for patient 2.

A visual comparison is presented in Figure 3, showing the same slice of the first and second MR image and the registration result. The manual and transformed delineations of the tumor are also shown.

## 4 Discussion and Conclusion

A nonrigid image registration algorithm was applied to intra-patient registration of MR and CT images recorded prior to and during radiotherapy treatment in two rectum and one head and neck patient.

Validation of nonrigid registration is a cumbersome task, as no ground truth is available. We choose to use manual delineations, as radiotherapists are familiar with delineation tasks. However, it is clear from the volume differences between the CT and MR delineations, that registration validation based on expert-guided segmentations may suffer from significant inaccuracies in the manual delineations. Therefore, it is difficult to judge whether the measured final error is due to an error during the delineation or an error during the registration [9].

Because soft-tissue structures are better visible in MR then in CT, we have presented the validation results obtained on MR images only. For comparison, Figure 3 shows a set of CT images and corresponding delineations before and after nonrigid registration. Visually the registration, including the tumor volume, seems to be acceptable. However, the manual delineation on the transformed image differs significantly.

The validation of the multitemporal registration of the MR images shows a significant improvement. This was especially clear in the rectum images. However, the initial error as well as the final error, were lower in the head-and-neck images . This difference is mainly caused by the higher complexity of the rectum region, due to the higher number of structures in the neighborhood of the tumor and the variation in bladder and rectum filling.

Future work will focus on the registration error in view of the clinical requirements imposed by radiotherapy. In clinical practice, not only the tumor but also a safety region around the tumor is irradiated, while trying to spare neighboring, possibly vital organs. Therefore, we will have to include the segmentation of more clinically relevant structures. Comparisons are possible between the treatment

plan, designed using the pre-treatment scan and delineations, and the manually and automatically delineated images recorded in between treatment sessions.

Also, multimodal registration of the MR and CT images is required. Preliminary results are very good for the head and neck patients, showing registration results approaching those of monomodal registration. Due to the large deformations caused by the difference in imaging position, accurate registration of the rectum patients is more difficult. Also, due to the uncertainty on the delineation, validation is still lacking.

Although our presented method can certainly be useful in current clinical practice, its use will be boosted with the availability of on-board cone beam CT images [10]. Registering these images to the planning CT will allow planning adaptation to the actual patient position on the table and at the time of treatment. Thus a more accurate dose delivery and possibly even smaller safety margins become possible. However, while current registrations can take up to 30 minutes, registration of on-board cone beam CT images should be performed in a few minutes or less.

## Acknowledgements

This research was sponsored in part by the Flemish Fund for Scientific Research (FWO-Vlaanderen, Belgium) under grant G.0566.06.

## References

1. van Herk, M., Bruce, A., Kroes, G., Shouman, T., Touw, A., Lebesque, J.: Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. International Journal of Radiation Oncology Biology Physics **33** (1995) 1311–1320
2. Schaly, B., Kempe, J., Bauman, G., Battista, J., Van Dyk, J.: Tracking the dose distribution in radiation therapy by accounting for variable anatomy. Physics in Medicine and Biology **49** (2004) 791–805
3. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Transactions om Medical Imaging **16** (1997) 187–198
4. Viola, P., Wells, W.M.: Alignment by Maximization of Mutual Information. Int. J. Comput. Vision **24** (1997) 137–154
5. Loeckx, D., Maes, F., Vandermeulen, D., Suetens, P.: Comparison between Parzen window interpolation and generalised partial volume estimation for non-rigid image registration using mutual information. In: Workshop on Biomedical Image Registration. (2006) accepted
6. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26** (1945) 297–302
7. Zijdenbos, A., Dawant, B., Margolin, R., Palmer, A.: Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Transactions om Medical Imaging **13** (1994) 716–724
8. Zou, K.H., Warfield, S.K., Bharatha, A., Tempany, C.M., Kaus, M.R., Haker, S.J., Wells, W.M., Jolesz, F.A., Kikinis, R.: Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index. Academic Radiology **11** (2004) 178–189

9. Collier, D.C., Burnett, S.S.C., Amin, M., Bilton, S., Brooks, C., Ryan, A., Roniger, D., Tran, D., Starkschall, G.: Assessment of consistency in contouring of normal-tissue anatomic structures. J Appl Clin Med Phys **4** (2003) 17–24
10. Oldman, M., Létourneau, D., Watt, L., Hugom, G., Yan, D., Lockman, D., Kim, L., Chen, P., Martinez, A., Wong, J.: Cone-beam-CT guided radiation therapy: A model for on-line application. Radiotherapy and Oncology **75** (2005) 271.e1–271.e8

# 3D Statistical Shape Modeling of Long Bones

Yuhui Yang[1,2], Anthony Bull[1], Daniel Rueckert[2], and Adam Hill[1]

[1] Department of Bioengineering
[2] Department of Computing
Imperial College London, United Kingdom

**Abstract.** The aims of this paper are to devise robust methods for the description of the variability in shapes of long bones using 3D statistical shape models (SSMs), and to test these on a dataset of humeri that demonstrate significant variability in shape. 30 primate humeri were CT scanned and manually segmented. SSMs were constructed from a training set of landmarks. The landmarks of the 3D shapes are extracted automatically using marching cubes and point correspondences are automatically obtained via a volumetric non-rigid registration technique using multiresolution B-Spline deformations. The surface registration resulted in no discernible differences between bone shapes, demonstrating the high accuracy of the registration method. An analysis of variations is applied on the shapes based on the model we built. The first mode of variation accounted for 42% of the variation in bone shape. This single component discriminated directly between great apes (including humans) and monkeys.

## 1 Introduction

Subject-specific preclinical surgical planning, intraoperative navigation, and musculoskeletal modelling require accurate definition of the shapes of bones. The most appropriate methods available require large imaging datasets and manual or semi-manual segmentation; these methods are not widely applicable and have significant health costs that includes exposure to ionizing radiation. Model based image analysis simplifying and stabilizing problems has been researched and widely used in computer vision, such as active shape models [5] and the active appearance models [4]. Statistical modelling is one of those techniques generating considerable interest for use in studying shape variations in anatomical shapes, as well as being taken as a basis for segmenting and interpreting images. The major difficulties in building statistical models are manual segmentation and determination of corresponding landmarks over the set of training shapes boundaries / surfaces. Finding automatic segmentation methods is not the primary aim of this work, thus only the process of landmark extraction will be discussed in this paper. The accuracy of the correspondences is of high importance because of the sensibility within shape parameterisation, which may lead to difficulty in defining shape constraints. Though manually labelling corresponding points on each sample is still being used in many applications, it is time-consuming and cumbersome for use in 2D images. It becomes impractical

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 306–314, 2006.

in 3D images due to the increasing number of landmarks required to describe the shape and the difficulties in identifying landmarks unambiguously. Numerous attempts have been made to automate the construction of 3D statistical shape models [6]. Kelemen et al [11] established surface correspondence by describing the surfaces by a series expansion using spherical harmonics while Wang et al [16] computed surface correspondence and triangulation in a hierarchical way after a geodesic interpolation to the initial sparse set of corresponding points generated using a shape-based matching procedure. Rhodri et al [8] described a model of automatically constructing optimal models by minimising the description length of the model. Currently, non-rigid registration is used by a number of authors to automate the 3D corresponding landmark set [3] [10] [9]. Brett and Taylor [3] constructed a mean shape from a set of smooth synthetic examples by the correspondence of their sparse polyhedral approximations and the production of a binary tree of merged shapes. Fleute and Lavalee [9] registered data sets using an elastic registration method provided by Szeliski and Lavalle based on octree-splines. Our method generating the statistical models is based on the work of Frangi et al [10].

## 2 Statistical Shape Models

The basic idea is to establish, from a training set, the pattern of 'legal' variation in the shapes and spatial relationships of structures in a given class of images. Statistical analysis is used to give an efficient parameterisation of this variability, providing a compact representation of shape [7]. SSMs are not equal to point distribution models (PDMs) because the statistics can be captured not only by the coordinates of feature points. The term 'landmark' can be lines (e.g. sutura, crista) or areas (e.g. pars, foramen), which can also describe the features a shape. For example, crest lines [19] [20] have a very strong anatomical significance and indicate salient structures so that they can be used as anatomical landmarks. In this paper, we consider landmarks as point features on the shape surface.

Given $n$ sample shapes $\{\boldsymbol{X}_i; i = 1, 2, \cdots, n\}$ in the same coordinate system, assuming that we have obtained $m$ 3D corresponding landmarks $\{(x_j, y_j, z_j); j = 1, 2, \cdots, m\}$ denoting the nodes of a surface triangulation for each shape, $\boldsymbol{X}_i$ can be represented by vectors $\boldsymbol{X}_a = (x_{a1}, y_{a1}, z_{a1}, x_{a2}, y_{a2}, z_{a2}, \cdots, x_{an}, y_{an}, z_{an})^T$ which form a distribution in a shape space. Principal component analysis (PCA) is used to reduce dimensionality by picking out the main axes of the distribution. Only the first few axes, which account for the majority of the variation, are modeled. The model can be represented as

$$\boldsymbol{X} = \bar{\boldsymbol{X}} + \boldsymbol{P}\boldsymbol{b} \tag{1}$$

$\bar{\boldsymbol{X}}$ is the mean vector allowing explicit measurement of the variation exhibited by each landmark. $\boldsymbol{b}$ is a $3n$ dimensional vector that controls shape variations by varying the values of it. $\boldsymbol{P}$ is a $3n$ by $t$ matrix whose columns are unit vectors along the principal axes of the distribution. $t$ is number of principal axes corresponding to the first largest eigenvalues. The shape generated is similar to those in the training set given $-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i}$.

## 3 Establishing Correspondences Using Surface Registration

Before establishing correspondences using non-rigid surface registration, all shapes are aligned using a rigid surface registration algorithm (ICP, [1]). Choosing a shape as the reference shape, all the other shapes are aligned into the coordinate system defined by the reference shape. Once all shapes have been translated and rotated into a common coordinate system, non-rigid registration is employed to establish correspondences among the training set. One of the most common mathematical models being using in non-rigid registration are splines (e.g. B-spline [17] [13] [14], thin-plate splines [2] [15] [12]). The registration algorithm employed here is a surface-based registration using multiresolution free-form deformations (FFD) which are modeled using B-splines. It is an extension of the image-based registration algorithm proposed by Rueckert et al. [17] and further developed by Schnabel et al. [18]. The basic idea of the algorithm is to embed a surface in a volumetric mesh which defines the continuous deformation field. The mesh is subsequently subdivided into higher resolution levels by inserting control points into the current level of control points and decreasing the mesh space; the mesh notes at each of the resolution levels are taken as the control points that are interpolated by a set of B-spline basis functions.

Let $\Omega = \{(x, y, z)|0 \leq x \leq X, 0 \leq y \leq Y, 0 \leq z \leq Z\}$ be the domain of the image, $\Phi$ be the $n_x \times n_y \times n_y$ mesh of control points $\phi_{i,j,k}$ at level $h$ with uniform spacing, the non-rigid transformation at level h can be written in the form of cubic B-spline function:

$$T^l_{local}(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(w) \phi_{i+l,j+m,k+n} \tag{2}$$

where $B_i$ is the $l$th uniform cubic B-spline basis function and $\phi$ denoting the control points. B-spline is locally controlled by changing the control points. The deformation is affected by the spacing of the control points: a larger spacing leads to a smoother global shape while a relatively smaller spacing leads to higher local deformation. A multiresolution FFD tends to generate a hierarchy of local deformation meshes by deforming a sequence of control points $\Phi^1, \Phi^2, \cdots, \Phi^H$. The local deformation is constructed by summing up all levels of the resolutions:

$$T_{local}(x, y, z) = \sum_{l=1}^{L} T^l_{local}(x, y, z) \tag{3}$$

To calculate the transformation T, the non-rigid surface registration algorithm minimizes the distance between every point in the reference surface and the closest point on the each of the other shapes. Once the closest points have been established for each point in the reference shape, the optimal control point values

of the FFM can be calculated using the algorithm proposed in [13]. This process is repeated until the distances between surfaces cannot be minimized any further. After registration, the reference shape is deformed to all other shapes. For each surface point the closest point on the other shapes is then assumed to be the corresponding point.

## 4 Results

A set of 30 primate humeri were CT scanned and manually segmented. There are maximum 500 slices for each humerus; the slice thickness is 1 mm; the image matrix dimension is 512 by 512; the in-plane resolution is 0.35 mm by 0.35 mm. All humeri were converted into 3D triangulated mesh surfaces by a marching cubes method. A decimated humerus surface and its triangular mesh are shown in figure 1. These 8 species of primates, including 3 cebuses, 4 colobuses, 3 gorillas, 4 macacas, 3 chimpanzees, 4 papios, 5 presbytises, and 4 human, encompassed all terrestrial locomotor types.

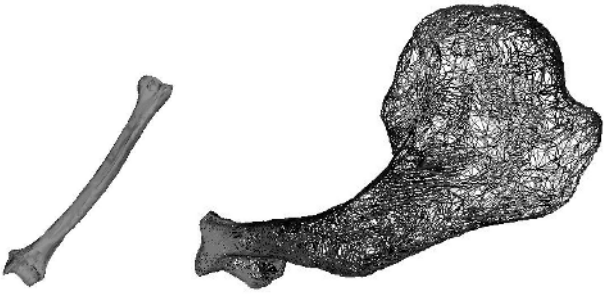
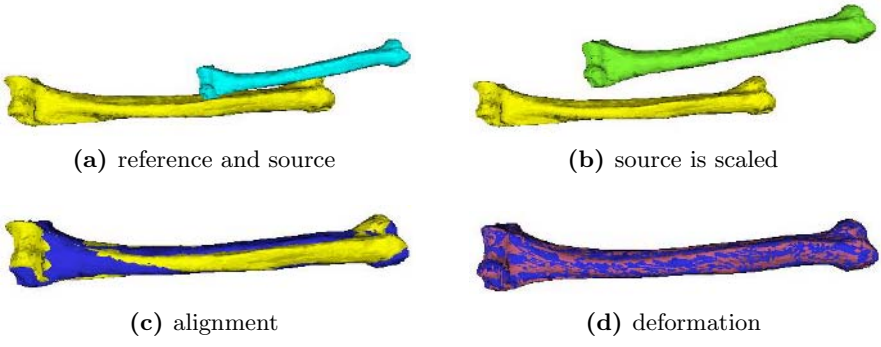


**Fig. 1.** A humerus surface and its triangular mesh

The 30 humeri are very different in size from specie to specie. For example, a gorilla humerus is about 5 times larger than a cebus humerus. At the current stage, we are only interested in shape variations, but not the size variations, therefore a bounding box scaling is applied to reduce complexity before the alignment of the dataset. The scaling factors for a specific humerus are defined as the difference of the longest axis between this humerus and a chosen reference humerus. If the reference humerus $\mathbf{T}$ is in a $\Delta X \times \Delta Y \times \Delta Z$ bounding box with $\Delta Z$ denoting the length of its longest axis, humerus $\mathbf{S}$ is in a $\Delta X_i \times \Delta Y_i \times \Delta Z_i$ bounding box with $\Delta Z_i$ denoting the length of its longest axis, $\mathbf{S}$ is scaled with scaling factors $t = \Delta Z/\Delta Z_i$ in $x$, $y$ and $z$ directions. The scaling procedure only changes the size of a humerus but does not change its shape.

In our dataset, a papio humerus with 13518 nodes is chosen as the reference shape. Applying scaling to each of the other 29 humeri in the dataset and aligning them to the coordinate system defined by the reference shape using rigid surface

registration, a new set of humeri in the same size is created at the same position and the same orientation. The reference humerus is then deformed to each of the other 29 humeri directly using multiresolution FFD. The deformation field is with four mesh subdivisions and its initial spacing is 40 mm. Figure 2 explains the pipeline of the above procedure. In Figure 2(d), two humeri are almost overlapping each other. Taking the landmarks on the deformed reference humerus as the dynamic atlas, the landmark set of a source humeri that have the minimum distance to the atlas are extracted.



**(a)** reference and source **(b)** source is scaled

**(c)** alignment **(d)** deformation

**Fig. 2.** Pipeline for finding correspondence

A training set with point correspondences is constructed for applying PCA. Figure 3(a) is a plot for the 30 eigenvalues and Figure 3(b) is a plot for their accumulated percentages. Over 90% of the variability in shape was defined by the first nine principle components. Figure 4 shows the first 3 modes of variation. The instances are generated by varying a single value in vector $b$ (equation (1)), fixing all others at zero standard deviations from the mean shape.

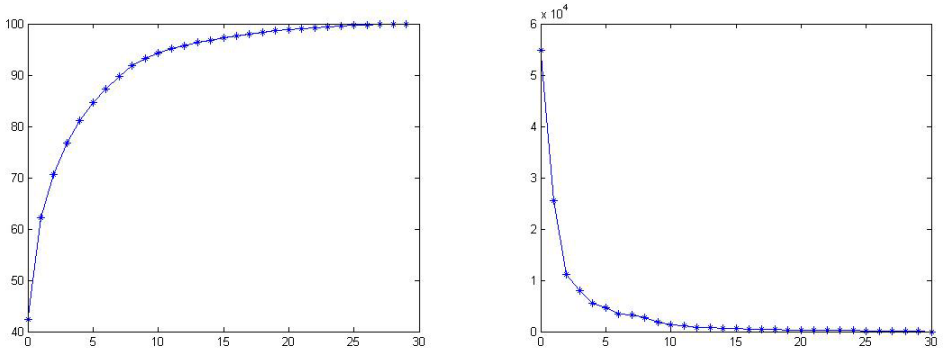



**Fig. 3.** Eigenvalues and their cumulative percentages

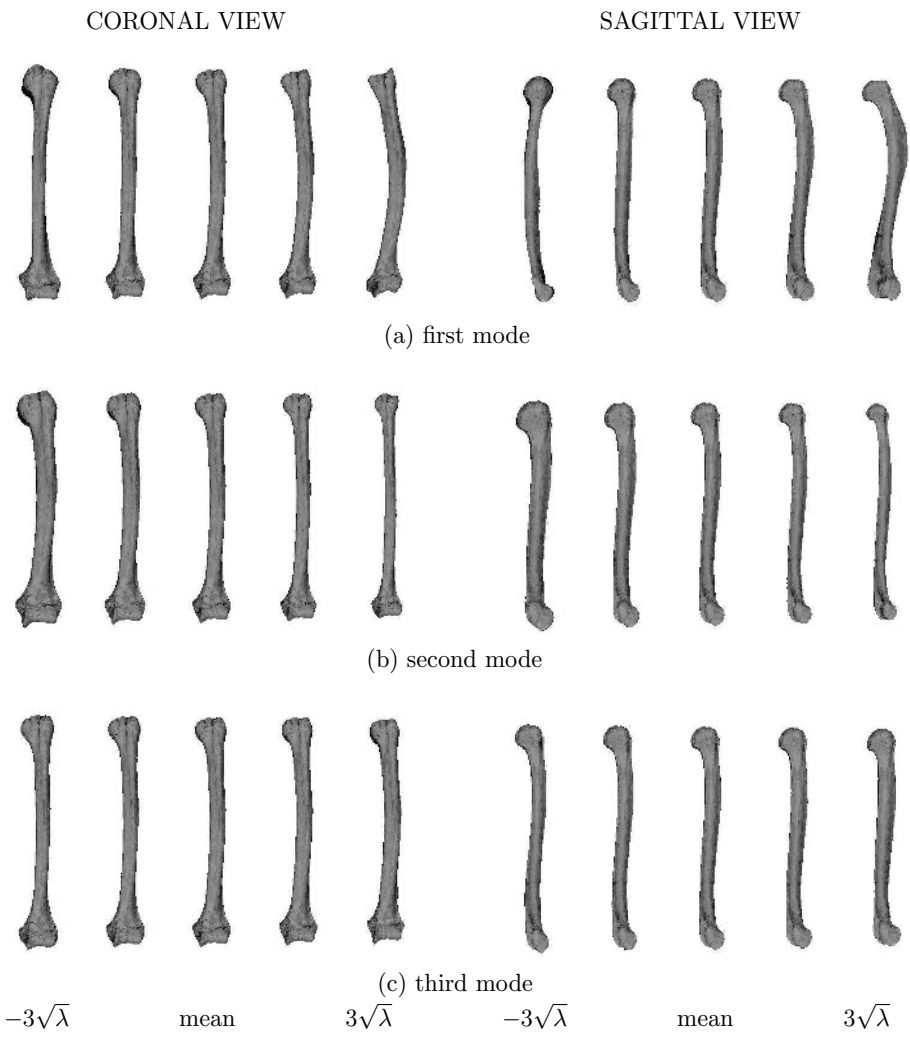



**Fig. 4.** First 3 modes in shape variations

Each specific humerus has its $\boldsymbol{b} = \Phi^{-1}(x - \bar{x})$ vector that distinguishes itself from other humeri. For each humerus, we take the first 3 values of its vector $\boldsymbol{b}$ as the coordinates for $x$, $y$ and $z$ axes; the shapes are represented by a 3D point. This resulted in two distinct clusters of which the primates in the left cluster are the great apes and humans, the primates in the right cluster are monkeys (Figure 5). The humeri are shown in sagittal view. The first values of the $\boldsymbol{b}$ vectors are the main features that distinguish the two clusters. The most obvious difference between these two clusters is bending structure which can be observed from the 1st mode of variations (Figure 4).

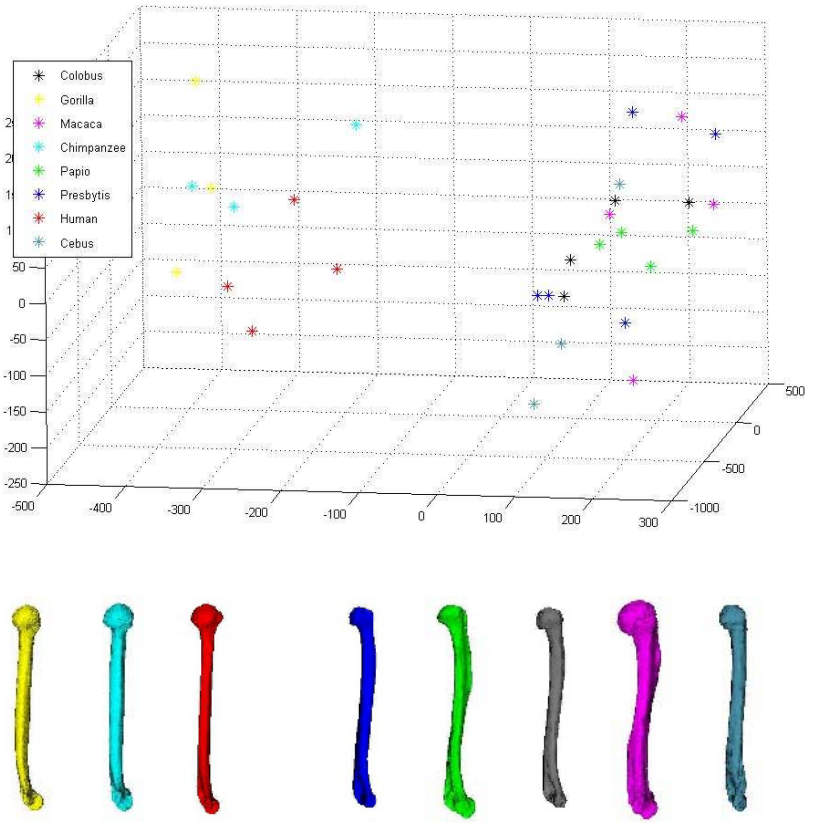



**Fig. 5.** Classifications (the humeri from the left to the right are gorilla, chimpanzee, human, presbytis, papio, colobus, macaca, cebus)

## 5 Conclusions

This paper has demonstrated a method for the automatic construction of 3D statistical shape models. The methodology is based on the automatic extraction of landmarks in a given class of images and the automatic establishment of correspondences among the landmark sets. The keys of establishing the correspondences are the generation of a dynamic atlas from the sample shapes and the deformations of the dynamic atlas to each of the other shapes.

The results show that the combination of a dynamic atlas and the multiresolution FFD non-rigid surface registration is robust to construct 3D SSMs for humeri from a variety species of primates. Since all the long bones consist of a

long shaft with two bulky ends or extremities, this procedure is a generic framework that can be extended to define the shapes of any other long bones (e.g. femur and tibia).

The results have also demonstrated that the shape information can be used as a classifier. The classification of our dataset shows that SSMs of bones can discriminate between bones that have different environmental influences that includes loading based on locomotor type, and taxonomy.

## References


1. P. Besl and N. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 1992.
2. Fred. L. Bookstein. Principal warps: Thin-plant splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), June 1989.
3. Alan D. Brett and Christopher J. Taylor. A method of automated landmark generation for automated 3d pdm construction. *Image Vision Comput*, 18(9):739 – 748, 2000.
4. T. F. Cootes, C. Beeston, G. J. Edwards, and C. J. Taylor. A unified framework for atlas matching using active appearance models. *Information Processing in Medical Imaging*, 1613:322 – 333, 1999.
5. T. F. Coots, D. Cooper, C. J. Taylor, and J. Graham. Active shape models their training and application. *Computer Vision and Image Understanding*, 61(1):38 – 59, January 1995.
6. Rh. H. Davies, Carole J. Twining, T. F. Cootes, and J. C. Taylor. 3d statistical shape models by direct optimisation of description length. *European Conference on Computer Vision ECCV 2*, 3:3 – 22, 2002.
7. Rhodri H. Davies, Carole J. Twining, Tim F. Cootes, John C. Waterton, and Chris J. Taylor. Automatic construction of optimal 3d statistical shape models. In *Medical Image Understanding and Analysis: BMVA*, pages 77 – 80, July 2002.
8. Rhodri H. Davies, Carole J. Twining, Timothy F. Cootes, John C. Waterton, and Christopher J. Taylor. A minimum description length approach to statistical shape modeling. *IEEE Transactions on Medical Imaging*, 21(5):525 – 537, 2002.
9. M. Fluete and S. Lavalee. Building a complete surface model from sparse data using statistical shape models: Application to computer assisted knee surgery. *MICCAI*, 8:878 – 887, 1998.
10. A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen. Automatic construction of multiple-object three dimensional statistical shape models: Application to cardiac modeling. *IEEE Transactions on Medical Imaging*, 21(9), September 2002.
11. A. Kelemen, G. Szekely, and G. Gerig. Elastic model-based segmentation of 3-d neuroradiological data sets. *IEEE Transactions on Medical Imaging*, 18(10):175 – 191, 1999.
12. R. J. A. Lapeer and R. W. Prager. 3d shape recovery of a newborn skull using thin-plate splines. *Computerized Medical Imaging and Graphics*, 24:193 – 204, 2000.
13. S. Lee, G. Wolberg, and S. Y. Shin. Scattered data interpolation with multilevel b-splines. *IEEE Transactions on Visulization and Computer Graphics*, 3:228 – 244, September 1997.
14. Steven Lobregt and Max A. Viergever. Discrete dynamic contour model. *IEEE Transactions on Medical Imaging*, 14(1), March 1995.

15. C. Lorenz and N. Krahnstover. Generation of point-based 3d statistical shape models for anatomical objects. *Computer Vision and Image Understanding*, 77:175 – 191, 2000.
16. K. T. Rajamani, L. Nolte, and M. Styner. Bone morphing with statistical shape models for enhanced visualization. *SPIE Medical Imaging*, 2004.
17. 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 Transaction on Medical Imaging*, 18:712 – 721, August 1999.
18. J. A. Schnabel, D. Rueckert, M. Quist, J. M. Blackall, A. D Castellano Smith, T. Hartkens, G. P. Penney, W. A. Hall, H. Liu, C. L. Truwit, F. A. Gerritsen, D. L. G. Hill, and D. J. Hawkes. A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations. *Medical Image Computing and Computer Assisted Intervention. ser. Lecture Notes in Computer Science. The Netherlands: Springer-Verlag*, 2208:573 – 581, 2001.
19. Gerard Subsol, Jean-Philippe Thirion, and Nicholas Ayache. A scheme for automatically building three-dimensional morphometric anatomical atlases: Application to a skull atlas. *Medical Image Analysis*, 2(1):37 – 60, 1998.
20. J. Ph. Thirion and A. Gourdon. The 3d marching lines algorithm. *Graphical Models Image Process Graphical Models and Image Processing*, 58(6):503 – 509, 1996.

# Point-Based Registration with Known Correspondence: Closed Form Optimal Solutions and Properties

Oskar Škrinjar*

Department of Biomedical Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA

**Abstract.** Point-based registration with known correspondence is often used as either a stand-alone method or a part of a more complex algorithm. The goal of this type of registration is to align two sets of points with the same number of corresponding points using a selected transformation type. Presented are closed form solutions for the transformation parameters that optimally align two point sets in the least squares sense for the following transformation types: rigid, similarity, rigid with nonuniform scales, and a linear combination of basis functions. It is shown that those registration methods whose underlying transformations form a group satisfy the identity, symmetry, transitivity, and distortion properties.

## 1 Introduction

This is a theoretical paper that provides a collection of the optimal (in the least squares) solutions for the point-based registration with known correspondence for the following transformation types: rigid, similarity, rigid with nonuniform scales, and linear combination of basis functions. The solutions for rigid [1] and similarity [2] cases have already been reported, and the original contribution of this paper are the solutions for the cases of rigid registration with nonuniform scales and a linear combination of basis functions, and the analysis of properties of the optimal solutions for each of the transformation types. Proofs of the results are omitted due to the limited space.

## 2 Notation and Definitions

### 2.1 Transformations, Registration Operators, Groups

Let $\mathbb{S}$ denote an $M$-dimensional metric space [3].

**Definition 1.** *An $M$-dimensional geometric transformation is a function* $\mathbf{T} : \mathbb{S} \mapsto \mathbb{S}$.

* Research supported by the NIH under grant number EB02957.

J.P.W. Pluim, B. Likar, and F.A. Gerritsen (Eds.): WBIR 2006, LNCS 4057, pp. 315–321, 2006.

Geometric transformations will be referred to as just transformations. Let $\mathcal{T}$ denote the set of transformations of certain type (e.g. affine transformations). Let $\mathbf{T}_{id}$ denote the identity transformation. Thus, $\mathbf{T}_{id}(\mathbf{r}) = \mathbf{r}, \quad \forall \mathbf{r} \in \mathbb{S}$. Let $\mathbb{P}$ denote the set of ordered sets of $N$ points from $\mathbb{S}$.

**Definition 2.** *A point-based registration operator is a function* $\mathbf{\Gamma} : \mathbb{P}^2 \mapsto \mathcal{T}$.

Point-based registration operators will be referred to as just registration operators. The fact that the elements of $\mathbb{P}$ are ordered sets of points makes this an operator for point-based registration with known correspondence (i.e. the $i$-th point from the first set of points corresponds to the $i$-th point from the second set of points).

**Definition 3.** *A set of transformations* $\mathcal{T}$ *with respect to transformation composition* $\circ$ *is said to form a group of transformations if it is closed* ($\forall \mathbf{T}_1, \mathbf{T}_2 \in \mathcal{T} \Rightarrow \mathbf{T}_1 \circ \mathbf{T}_2 \in \mathcal{T}$)*, if it contains the identity transformation* ($\mathbf{T}_{id} \in \mathcal{T}$) *and if every transformation in* $\mathcal{T}$ *has an inverse in* $\mathcal{T}$ ($\forall \mathbf{T} \in \mathcal{T}$ *there is* $\mathbf{T}^{-1} \in \mathcal{T}$ *such that* $\mathbf{T} \circ \mathbf{T}^{-1} = \mathbf{T}_{id}$)*.*

Note that the standard group definition [3] requires associativity, which is trivially satisfied for the composition of transformations ($\forall \mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3 \in \mathcal{T} \Rightarrow (\mathbf{T}_1 \circ \mathbf{T}_2) \circ \mathbf{T}_3 = \mathbf{T}_1 \circ (\mathbf{T}_2 \circ \mathbf{T}_3)$) and is therefore not included in Def. 3.

### 2.2 Transformation Types

Let $I$ denote an $M \times M$ identity matrix.

**Definition 4.** *An* $M$*-dimensional rotation matrix is an* $M \times M$ *orthonormal matrix* $R$ ($RR^T = I$) *for which* $\det(R) = 1$.

**Definition 5.** *An* $M$*-dimensional rigid transformation is a transformation of the type*

$$\mathbf{T}(\mathbf{r}) = R\mathbf{r} + \mathbf{t},$$

*where* $R$ *is an* $M$*-dimensional rotation matrix and* $\mathbf{t}$ *an* $M \times 1$ *translation vector.*

An $M$-dimensional rigid transformation has $\frac{M(M+1)}{2}$ degrees of freedom.

**Definition 6.** *An* $M$*-dimensional similarity transformation is a transformation of the type*

$$\mathbf{T}(\mathbf{r}) = sR\mathbf{r} + \mathbf{t},$$

*where scalar* $s$ *is a uniform scale factor,* $R$ *is an* $M$*-dimensional rotation matrix and* $\mathbf{t}$ *an* $M \times 1$ *translation vector.*

An $M$-dimensional similarity transformation has $\frac{M(M+1)}{2}+1$ degrees of freedom.

**Definition 7.** *An* $M$*-dimensional rigid transformation with nonuniform scales is a transformation of the type*

$$\mathbf{T}(\mathbf{r}) = SR\mathbf{r} + \mathbf{t},$$

*where* $S$ *is an* $M \times M$ *diagonal matrix,* $R$ *is an* $M$*-dimensional rotation matrix and* $\mathbf{t}$ *an* $M \times 1$ *translation vector. Matrix* $S$ *is called the scale matrix and its diagonal elements are called nonuniform scale factors.*

An $M$-dimensional rigid transformation with nonuniform scales has $\frac{M(M+3)}{2}$ degrees of freedom.

**Definition 8.** *An $M$-dimensional affine transformation is a transformation of the type*

$$\mathbf{T}(\mathbf{r}) = A\mathbf{r} + \mathbf{t},$$

*where $A$ is an $M \times M$ matrix and $\mathbf{t}$ an $M \times 1$ translation vector.*

An $M$-dimensional affine transformation has $M(M+1)$ degrees of freedom.

**Definition 9.** *An $M$-dimensional linear combination of basis functions is a transformation of the type*

$$\mathbf{T}(\mathbf{r}) = \sum_{k=1}^{K} \mathbf{w}_k B_k(\mathbf{r}),$$

*where $B_k$ are basis functions and $M \times 1$ vectors $\mathbf{w}_k$ are the transformation parameters.*

This transformation representation has $KM$ degrees of freedom. Linear combination of basis functions is a general transformation model that can represent a wide class of transformations including affine, Fourier series [4], thin plate splines (with fixed nodes) [5], and polynomial transformations of any order (e.g. B-splines [6]). However, it does not include rigid, similarity and rigid with nonuniform scales transformations.

Note that the rigid, similarity (with a nonzero uniform scale factor), rigid with nonuniform scales (with nonzero nonuniform scale factors) and (non-singular) affine transformations form transformation groups while the linear combination of basis functions in general does not.

### 2.3 Properties

A registration operator needs to satisfy the following properties, which are given both descriptively and formally.

**Identity.** A registration operator, when applied to two identical point sets should generate the identity transformation. Formally,

$$\mathbf{\Gamma}(p, p) = \mathbf{T}_{id}.$$

**Symmetry.** When it is applied to two point sets, the obtained transformation should be the inverse of the transformation obtained when the order of the point sets is reversed. Formally,

$$\mathbf{\Gamma}(p, q) = [\mathbf{\Gamma}(q, p)]^{-1}.$$

**Transitivity.** The operator should be transitive, i.e. for any three point sets, the generated transformation from the second to the third set composed with the generated transformation from the first to the second set should be equal to the generated transformation from the first to the third set. Formally,

$$\mathbf{\Gamma}(q, r) \circ \mathbf{\Gamma}(p, q) = \mathbf{\Gamma}(p, r).$$

**Distortion.** If one of the point sets is geometrically deformed (distorted), the obtained transformation should reflect the distortion. The distortion functions should be from the set of transformations in which the solution is sought. Formally,

$$\mathbf{\Gamma}(\mathbf{T}(p), \mathbf{U}(q)) = \mathbf{U}^{-1} \circ \mathbf{\Gamma}(p, q) \circ \mathbf{T}.$$

Here $\mathbf{T}(p)$ represent the point set obtained by applying transformation $\mathbf{T}$ to each point from point set $p$. The above properties have to hold $\forall p, q, r \in \mathbb{P}$ and $\forall \mathbf{T}, \mathbf{U} \in \mathcal{T}$. These properties can be simplified without loss of generality to the following list.

$$\textbf{(i)}\ \mathbf{\Gamma}(p \circ \mathbf{T}(p)) = \mathbf{T}, \quad \forall \mathbf{T} \in \mathcal{T}, \quad \forall p \in \mathbb{P}$$

$$\textbf{(ii)}\ \mathbf{\Gamma}(p, q) = [\mathbf{\Gamma}(q, p)]^{-1}, \quad \forall p, q \in \mathbb{P}$$

$$\textbf{(iii)}\ \mathbf{\Gamma}(q, r) \circ \mathbf{\Gamma}(p, q) = \mathbf{\Gamma}(p, r), \quad \forall p, q, r \in \mathbb{P}$$

It is not difficult to prove that this set of properties is equivalent to the first one. For this reason, only the set of properties **(i)** - **(iii)**, being the simpler of the two, will be used from now on.

The above properties can be defined only for the registration operators that are defined on those sets of transformation $\mathcal{T}$ that form a transformation group. E.g. B-splines do not form a transformation group because (among other reasons) the inverse of a B-spline transformation is not a B-spline transformation. This prevents one to define the symmetry property (and others) with $\mathcal{T}$ being B-spline transformations.

### 2.4 Problem Statement

The goal is to provide a closed form solution for the transformation parameters that optimally align two points sets with known correspondence. The optimal alignment in the least squares sense is defined here.

**Definition 10.** *Let* $\mathbf{T}_{\alpha_1,\ldots,\alpha_K}$ *represent a transformation model with* $K$ *parameters* $\alpha_i$*. The parameters that optimally align two corresponding points sets* $\{\mathbf{x}_1, \ldots \mathbf{x}_N\}$ *and* $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$ *minimize the sum of squared distances*

$$\sum_{i=1}^{N} \left\| \mathbf{T}_{\alpha_1,\ldots,\alpha_K}(\mathbf{x}_i) - \mathbf{y}_i \right\|^2.$$

## 3 Optimal Transformations

### 3.1 Transformations Involving Rotation Matrix

The optimal parameters for the transformations involving rotation matrix need to satisfy the orthonormality of the rotation matrix which makes these cases distinct from the case of linear combination of basis functions. The computation of the optimal translation vector for these cases is based on the following lemma.

**Lemma 1.** *The optimal rigid transformation with nonuniform scales between two sets of corresponding points maps the centroid of the first set of points to the centroid of the second set of points.*

This lemma is a generalization of the corresponding result for the case of rigid [1] and similarity [2] transformation. If the two sets of points are $\{\mathbf{x}_1, \ldots \mathbf{x}_N\}$ and $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$ and if the optimal scale matrix $S$ and rotation matrix $R$ are known, then it follows from Lemma 1 that the optimal translation vector is

$$\mathbf{t} = \frac{1}{N}\sum_i \mathbf{y}_i - SR\frac{1}{N}\sum_i \mathbf{x}_i. \tag{1}$$

This allows one to eliminate the translation from the problem by de-meaning the points, determine the optimal rotation matrix (and uniform scale factor or nonuniform scale factors) and then determine the optimal translation vector from Eq. 1. Let the de-meaned points be

$$\mathbf{X}_i = \mathbf{x}_i - \frac{1}{N}\sum_i \mathbf{x}_i \quad \text{and} \quad \mathbf{Y}_i = \mathbf{y}_i - \frac{1}{N}\sum_i \mathbf{y}_i.$$

**Rigid Transformation** Let

$$\Sigma = \frac{1}{N}\sum_i \mathbf{Y}_i \mathbf{X}_i^T.$$

Let the singular value decomposition of $\Sigma$ be $UDV^T$, and

$$C = \begin{cases} I & \text{if} \quad \det(\Sigma) \geq 0 \\ diag(1, 1, \ldots, 1, -1) & \text{if} \quad \det(\Sigma) < 0 \end{cases}$$

In the case of rigid registration, the optimal rotation matrix is

$$R = UCV^T. \tag{2}$$

This is a result from [2] and is given here for completeness purposes. Once the optimal $R$ is known, the optimal translation vector is obtained from Eq. 1 using $S = I$ (there is no scaling present in the case of rigid transformation).

**Similarity Transformation.** In the case of similarity transformation, the optimal rotation matrix is given by Eq. 2, the optimal uniform scaling factor is

$$s = \frac{N}{\sum_i \mathbf{X}_i^T \mathbf{X}_i}\text{tr}(DC),$$

and the optimal translation vector is obtained from Eq. 1 using $S = sI$. Again, this is a result from [2] and is given here for completeness purposes.

**Rigid Transformation with Nonuniform Scales**

**Lemma 2.** *Let* $\mathbf{X}_i = \begin{bmatrix} X_{i1} \cdots X_{iM} \end{bmatrix}^T$, $\mathbf{Y}_i = \begin{bmatrix} Y_{i1} \cdots Y_{iM} \end{bmatrix}^T$, $\alpha_j = \sum_i Y_{ij}\mathbf{X}_i$, *and* $\beta_j = \sum_i X_{ij}^2$. *Let the optimal rotation matrix for the rigid transformation with nonuniform scales be* $R = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_M \end{bmatrix}$, *where* $\mathbf{r}_i$ *are the rows of* $R$. *Then the optimal nonuniform scale factors are*

$$s_k = \frac{\mathbf{r}_k \alpha_k}{\beta_k}, \quad k = 1, \ldots, M, \tag{3}$$

*and the corresponding sum of squared distances is*

$$\sum_i \mathbf{Y}_i^T \mathbf{Y}_i - \sum_j \frac{(\mathbf{r}_j \alpha_j)^2}{\beta_j}. \tag{4}$$

The purpose of Lemma 1 is to remove the translation vector from the problem and the purpose of Lemma 2 is to remove the nonuniform scale factors from the problem. The idea is to find the rotation matrix by minimizing Eq. 4, then compute the nonuniform scale factors from Eq. 3, and finally use Eq. 1 to compute the translation vector. The minimization of Eq. 4 needs to be done with the constrains that $R$ is orthonormal and that $\det(R) = 1$. One way to solve for $R$ is to use a parametric representation for $R$ and then minimize Eq. 4 with respect to the parameters (angles). E.g. in the 2D case

$$R = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix},$$

and the optimal angle of rotation is

$$\phi = \frac{1}{2} \tan^{-1} \frac{k_3}{k_1 - k_2},$$

where

$$k_1 = \frac{\alpha_{11}^2}{\beta_1} + \frac{\alpha_{22}^2}{\beta_2}, \quad k_2 = \frac{\alpha_{12}^2}{\beta_1} + \frac{\alpha_{21}^2}{\beta_2}, \quad \text{and} \quad k_3 = 2\frac{\alpha_{21}\alpha_{22}}{\beta_2} - 2\frac{\alpha_{11}\alpha_{12}}{\beta_1}.$$

### 3.2 Linear Combination of Basis Functions

**Theorem 1.** *The transformation* $\mathbf{T}(\mathbf{r}) = \sum_{k=1}^K \mathbf{w}_k B_k(\mathbf{r})$ *that optimally (in the least squares sense) aligns two sets of corresponding points* $\{\mathbf{x}_1, \ldots \mathbf{x}_N\}$ *and* $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$ *has the following parameters*

$$\begin{bmatrix} \mathbf{w}_1 \cdots \mathbf{w}_K \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 \cdots \mathbf{y}_N \end{bmatrix} A (A^T A)^{-1},$$

*where*

$$A = \begin{bmatrix} B_1(\mathbf{x}_1) & \cdots & B_K(\mathbf{x}_1) \\ \vdots & & \vdots \\ B_1(\mathbf{x}_N) & \cdots & B_K(\mathbf{x}_N) \end{bmatrix}.$$

## 4 Properties of Optimal Registration Operators

Let $\mathbf{\Gamma}_R$ represent the optimal rigid registration operator, $\mathbf{\Gamma}_S$ represent the optimal similarity registration operator, $\mathbf{\Gamma}_N$ represent the optimal rigid with nonuniform scales registration operator, and $\mathbf{\Gamma}_A$ represent the optimal affine registration operator.

**Theorem 2.** *Registration operators* $\mathbf{\Gamma}_R$, $\mathbf{\Gamma}_S$, $\mathbf{\Gamma}_N$, $\mathbf{\Gamma}_A$ *satisfy identity, symmetry, transitivity, and distortion properties.*

## 5 Summary

Presented is a collection of the optimal (in the least squares sense) solutions for the point-based registration with known correspondence for the following transformation types: rigid, similarity, rigid with nonuniform scales, and linear combination of basis functions. It is shown that those registration methods whose underlying transformations form a group satisfy the identity, symmetry, transitivity, and distortion properties. While proofs are not included in this paper, the optimal solutions and the properties can be proven mathematically, which means than they hold for any sets of points.

## References

1. Arun, K., Huang, T.S., Blostein, S.D.: Least-squares fitting of two 3-d point sets. IEEE Transactions on Pattern Analysis and Machine Intelligence **9** (1987) 698–700
2. Umeyama, S.: Least-squares estimation of transformation parameters between two point sets. IEEE Transactions on Pattern Analysis and Machine Intelligence **13** (1991) 376–380
3. Bronshtein, I.N., Semendyayev, K.A.: Handbook of Mathematics. Third edn. Springer-Verlag, Berlin (1998)
4. Christensen, G.E., Johnson, H.J.: Consistent image registration. IEEE Transactions on Medical Imaging **20** (2001) 568–582
5. Bookstein, F.L.: Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence **11** (1989) 567–585
6. Rueckert, D., et al.: Nonrigid registration using free-form deformations: applications to breast mr images. IEEE Transactions on Medical Imaging **18** (1999) 712–721

# Author Index